PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A01K 67/00, A61K 38/09, C12N 15/00, 15/16, C07H 21/04, C07K 7/06, 14/575, 16/18, 16/26, C12P 21/06, 21/08, G01N 33/53

A1

(11) International Publication Number: WO 95/12309

(43) International Publication Date: 11 May 1995 (11.05.95)

(21) International Application Number: PCT/US94/12763

(22) International Filing Date: 4 November 1994 (04.11.94)

(30) Priority Data:
08/147,771 5 November 1993 (05.11.93) US

(71) Applicants *(for all designated States except US)*: THE BOARD OF TRUSTEES OF LELAND STANFORD, JR. UNIV [US/US]; ERSITY:, Suite 350, 900 Welch Road, Palo Alto, CA 94304-1858 (US). STATE OF OREGON, acting by and through THE OREGON STATES BOARD OF HIGHER EDUCATION on behalf of THE OREGON HEALTH SCIENCES UNIVERSITY [US/US]; 3181 S.W. Sam Jackson Park Road, Portland, OR 97201 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: FERNALD, Russell, D. [US/US]; 718 Alvarado Row, Stanford, CA 97305 (US). ADELMAN, John, P. [US/US]; 2433 S.W. Mitchell Street, Portland, OR 97201 (US).

(74) Agents: KRUSE, Norman, J. et al.; Townsend and Townsend Khourie and Crew, Steuart Street Tower, 20th floor, One Market Plaza, San Francisco, CA 94105-1492 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.

(54) Title: NUCLEIC ACIDS ENCODING [His^5, Trp^7, Tyr^8]-GnRH PREPROHORMONE AND [Ser^8]-GnRH PREPROHORMONE AND THEIR USES

(57) Abstract

This invention relates to DNA and protein compositions useful, for instance, in the regulation of reproductive function, and which are also useful in the development of transgenic animals with desirable reproductive characteristics. Specifically, this invention relates to DNA and protein compositions for the precursor protein for the chicken II form of gonadotropin-releasing hormone and to the [Ser^8]-GnRH preprohormone.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

Nucleic Acids Encoding [His^5,Trp^7,Tyr^8]-GnRH Preprohormone and [Ser^8]-GnRH Preprohormone and Their Uses

BACKGROUND OF THE INVENTION

This invention relates to DNA and protein compositions useful, for instance, in the regulation of reproductive function, and which are also useful in the development of transgenic animals with desirable reproductive characteristics. Specifically, this invention relates to DNA and protein compositions for the precursor protein for the chicken II form of gonadotropin-releasing hormone and to the [Ser^8]-GnRH preprohormone.

Gonadotropin-releasing hormone(GnRH) is an important reproductive hormone in vertebrates. GnRH is a neuropeptide which regulates the production of follicle stimulating hormone (FSH) and luteinizing hormone (LH). In mammals, GnRH is produced by the hypothalamus and regulates the release of FSH and LH, which in turn regulate the development and function of the reproductive organs.

There are eight different forms of GnRH that have been identified in different vertebrate species (see figure 1.) For a review of the species distribution and properties of the different forms of GnRH, *See* Sherwood, N.M., *et al.* (1993) *Endocrine Rev.* 14:241-254. For instance, the more recently evolved mammals have a form of GnRH designated mammalian GnRH. A number of primitive placental mammals, nonplacental mammals and other vertebrates have more than one form of GnRH. One form of GnRH that is particularly widespread in this latter group of animals and has been named chicken II GnRH, since it was the second form of GnRH that was discovered in the chicken. This form of GnRH is also known as [His^5,Trp^7,Tyr^8]-GnRH since it has amino acid substitutions in the 5, 7 and 8 positions as compared to mammalian GnRH (see figure 1).

[His^5,Trp^7,Tyr^8]-GnRH is a potent stimulator of reproductive function in a wide variety of species. It is particularly active in fish. [His^5,Trp^7,Tyr^8]-GnRH is the most potent form of GnRH in fish. (*See* Sherwood, N.M., *et al.*, *supra* and Ngamvongchon, S., *et al.* (1992) *Regulatory Peptides* 42:63-73.) [His^5,Trp^7,Tyr^8]-GnRH also apparently has significant biological activity in more recently evolved mammals.

One of the problems in fish aquaculture is control of fish reproduction. For some species of fish, it is difficult to induce the fish to reproduce in captivity. Alternatively, it may be difficult to induce fish in captivity to reproduce at the time of year that is normal to the species. One approach to the control of reproduction in fish for aquaculture is the development of transgenic fish by introduction of DNA encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone and the [Ser^8]-GnRH preprohormone.

Many of the potential uses of GnRH, for example fish aquaculture, require identification and isolation of the gene encoding the precursor protein from which the peptide is produced. The sequence of the [His^5,Trp^7,Tyr^8]-GnRH and the [Ser^8]-GnRH preprohormones and the sequence of the genes which encode these proteins have not been described in the prior art. Identification of the genes encoding these sequences will facilitate production and use of the decapeptide hormones and other biologically active fragments in a variety of applications. These and other needs are addressed by the present invention.

SUMMARY OF THE INVENTION

The present invention provides compositions for isolated nucleic acids encoding vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormones and [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptides. The invention also provides for isolated nucleic acids encoding vertebrate [Ser^8]-GnRH preprohormones and [Ser^8]-GnRH preprohormone GAP peptides. In addition, the invention also provides for compositions comprising the above

preprohormone polypeptides. The preprohormone polypeptides may be, for example, recombinantly produced.

Nucleic acid probes that selectively hybridize to nucleic acids encoding the above preprohormone polypeptides are also provided. The invention provides methods of detecting nucleic acids encoding the preprohormone polypeptides by nucleic hybridization assays utilizing these probes. Antibodies that are reactive with the preprohormone polypeptides are provided, along with immunoassay methods based on these antibodies and which detect the preprohormone polypeptides. These nucleic acid hybridization assays and immunoassays are useful as measurements of reproductive function in a variety of different situations.

The invention further provides for transgenic animals created by the introduction of a DNA construct encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone or a [Ser^8]-GnRH preprohormone. These transgenic animals include fish species useful in aquaculture and farm animal species, such as chickens and other animals.

DEFINITIONS

Abbreviations for the twenty naturally occurring amino acids follow conventional usage. In the polypeptide notation used herein, the left-hand direction is the amino terminal direction and the right-hand direction is the carboxy-terminal direction, in accordance with standard usage and convention. Similarly, unless specified otherwise, the left hand end of single-stranded polynucleotide sequences is the 5' end; the left hand direction of double-stranded polynucleotide sequences is referred to as the 5' direction. The direction of 5' to 3' addition of nascent RNA transcripts is referred to as the transcription direction; sequence regions on the DNA strand having the same sequence as the RNA and which are 5' to the 5' end of the RNA transcript are referred to as "upstream sequences"; sequence regions on the DNA strand having the same sequence as the RNA and which are 3' to the 3' end of the RNA transcript are referred to as "downstream sequences".

The nucleotide codes shown below are used in the nucleic acid sequences of the invention. This code has been adopted by the IUPAC-IUB Biochemical Nomenclature Commission.

Code	Group	Nucleotide(s)
A	A	adenine
C	C	cytosine
G	G	guanine
T	T	thymine
U	U	uracil
Y	C or T(U)	pyrimidine
R	A or G	purine
M	A or C	amino
K	G or T(U)	keto
S	G or C	3 hydrogen bonds
W	A or T(U)	2 hydrogen bonds
H	A or C or T(U)	not-G
B	G or T(U) or C	not-A
V	G or C or A	not-T(U)
D	G or A or T(U)	not-C
N	G,A,C or T(U)	any

The term "nucleic acids", as used herein, refers to either DNA or RNA. "Nucleic acid sequence" or "polynucleotide sequence" refers to a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. It includes both self-replicating plasmids, infectious polymers of DNA or RNA and nonfunctional DNA or RNA.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity", and "substantial identity". A "reference sequence" is a defined

sequence used as a basis for a sequence comparison; a reference sequence may be a subset of a larger sequence, for example, as a segment of a full-length cDNA or gene sequence given in a sequence listing, such as the nucleic acid sequence of figure 2, or may comprise a complete cDNA or gene sequence. Generally, a reference sequence is at least 20 nucleotides in length, frequently at least 25 nucleotides in length, and often at least 50 nucleotides in length. Since two polynucleotides may each (1) comprise a sequence (*i.e.*, a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) may further comprise a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity.

A "comparison window", as used herein, refers to a conceptual segment of at least 20 contiguous nucleotide positions wherein a polynucleotide sequence may be compared to a reference sequence of at least 20 contiguous nucleotides and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences.

Optimal alignment of sequences for aligning a comparison window may be conducted by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. (USA)* 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection, and the best alignment (*i.e.*, resulting in the highest percentage of sequence similarity over the comparison window) generated by the various methods is selected. Preferred programs for

this use include GAP, BESTFIT AND FASTA, with standard parameters.

The term "sequence identity" means that two polynucleotide sequences are identical (*i.e.*, on a nucleotide-by-nucleotide basis) over the window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

The terms "substantial identity" or "substantial sequence identity" as applied to nucleic acid sequences and as used herein and denote a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 85 percent sequence identity, preferably at least 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison window of at least 20 nucleotide positions, frequently over a window of at least 25-50 nucleotides, wherein the percentage of sequence identity is calculated by comparing the reference sequence to the polynucleotide sequence which may include deletions or additions which total 20 percent or less of the reference sequence over the window of comparison. The reference sequence may be a subset of a larger sequence, for example, as a segment of the full-length [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH sequences disclosed herein.

As applied to polypeptides, the terms "substantial identity" or "substantial sequence identity" mean that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least 80 percent sequence identity, preferably at least 90 percent sequence identity, more preferably at least 95 percent sequence identity or more (*e.g.*, 99 percent sequence

identity). Preferably, residue positions which are not identical differ by conservative amino acid substitutions. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

The term "isolated" or "substantially pure" means a compound is the predominant species present (*i.e.*, on a molar basis it is more abundant than any other individual species in the composition), and preferably a substantially purified fraction is a composition wherein the object species comprises at least about 50 percent (on a molar basis) of all macromolecular species present. Generally, a substantially pure composition will comprise more than about 80 to 90 percent of all macromolecular species present in the composition. Most preferably, the species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single macromolecular species. For example, "isolated" or "substantially pure", when referring to nucleic acids, refers to those that have been purified away from other chromosomal or extrachromosomal DNA or RNA by standard techniques, including alkaline/SDS treatment, CsCl banding, column chromatography, and other techniques well known in the art. *See*, F. Ausubel, *et al.*, ed. *Current Protocols in Molecular*

Biology, Greene Publishing and Wiley-Interscience, New York (1987), incorporated herein by reference.

"Nucleic acid probes" may be DNA or RNA fragments. DNA fragments prepared, for example, by digesting plasmid DNA, or by use of PCR, or synthesized by either the phosphoramidite method described by Beaucage and Carruthers, *Tetrahedron Lett.* 22:1859-1862 (1981), or by the triester method according to Matteucci, *et al.*, *J. Am. Chem. Soc.*, 103:3185 (1981), both incorporated herein by reference. A double stranded fragment may then be obtained, if desired, by annealing the chemically synthesized single strands together under appropriate conditions or by synthesizing the complementary strand using DNA polymerase with an appropriate primer sequence. Where a specific sequence for a nucleic acid probe is given, it is understood that the complementary strand is also identified and included. The complementary strand will work equally well in situations where the target is a double-stranded nucleic acid.

The phrase "selectively hybridizing to" refers to a nucleic acid probe that hybridizes, duplexes or binds only to a particular target DNA or RNA sequence when the target sequences are present in a preparation of total cellular DNA or RNA. "Complementary" or "target" nucleic acid sequences refer to those nucleic acid sequences which selectively hybridize to a nucleic acid probe. Proper annealing conditions depend, for example, upon a probe's length, base composition, and the number of mismatches and their position on the probe, and must often be determined empirically. For discussions of nucleic acid probe design and annealing conditions, see, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989) or *Current Protocols in Molecular Biology*, F. Ausubel *et al.*, ed. Greene Publishing and Wiley-Interscience, New York (1987).

The phrase "nucleic acid sequence encoding" refers to a nucleic acid which directs the expression of a specific protein or peptide. The nucleic acid sequences include both the DNA strand sequence that is transcribed into RNA and the

RNA sequence that is translated into protein. The nucleic acid sequences include both the full length nucleic acid sequences as well as non-full length sequences derived from the full length protein. It being further understood that the sequence includes the degenerate codons of the native sequence or sequences which may be introduced to provide codon preference in a specific host cell.

The phrase "expression cassette", refers to nucleotide sequences which are capable of affecting expression of a structural gene in hosts compatible with such sequences. Such cassettes include at least promoters and optionally, transcription termination signals. Additional factors necessary or helpful in effecting expression may also be used as described herein.

The term "operably linked" as used herein refers to linkage of a promoter upstream from a DNA sequence such that the promoter mediates transcription of the DNA sequence.

The term "vector", refers to viral expression systems, autonomous self-replicating circular DNA (plasmids), and includes both expression and nonexpression plasmids. Where a recombinant microorganism or cell culture is described as hosting an "expression vector," this includes both extrachromosomal circular DNA and DNA that has been incorporated into the host chromosome(s). Where a vector is being maintained by a host cell, the vector may either be stably replicated by the cells during mitosis as an autonomous structure, or is incorporated within the host's genome.

The term "plasmid" refers to an autonomous circular DNA molecule capable of replication in a cell, and includes both the expression and nonexpression types. Where a recombinant microorganism or cell culture is described as hosting an "expression plasmid", this includes both extrachromosomal circular DNA molecules and DNA that has been incorporated into the host chromosome(s). Where a plasmid is being maintained by a host cell, the plasmid is either being stably replicated by the cells during mitosis as an autonomous structure or is incorporated within the host's genome.

The phrase "recombinant protein" or "recombinantly produced protein" refers to a peptide or protein produced using non-native cells that do not have an endogenous copy of DNA able to express the protein. The cells produce the protein because they have been genetically altered by the introduction of the appropriate nucleic acid sequence. The recombinant protein will not be found in association with proteins and other subcellular components normally associated with the cells producing the protein.

"Specifically immunoreactive" refers to a binding reaction between an antibody and antigen which is determinative of the presence of the antigen in the presence of a heterogeneous population of proteins and other biological macromolecules or in a biological sample. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone with the amino acid sequence depicted in Seq. ID No. 2 or to the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone sequence depicted in Seq. ID No. 18 can be selected to obtain antibodies specifically immunoreactive with [His^5,Trp^7,Tyr^8]-GnRH preprohormone proteins and not to other proteins. Homologous proteins to the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone and the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone encompass [His^5,Trp^7,Tyr^8]-GnRH preprohormone species, but do not include other proteins. Similarly, antibodies raised to the *H. burtoni* [Ser^8]-GnRH preprohormone with the amino acid sequence depicted in Seq. ID No. 20 can be selected to obtain antibodies specifically immunoreactive with [Ser^8]-GnRH preprohormone proteins and not to other proteins. Homologous proteins to the *H. burtoni* [Ser^8]-GnRH preprohormone encompass [Ser^8]-GnRH preprohormone species, but do not include other proteins.

A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

"Biological sample" as used herein refers to any sample obtained from a living organism or from an organism that has died. Examples of biological samples include body fluids and tissue specimens.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A is a comparison of known GnRH amino acid sequences. The primary structures of the eight known GnRH forms and of the first nine amino-terminal residues from yeast α-mating factor are shown.

Fig. 1B shows the predicted amino acid sequence of the preprohormone for *H. burtoni* [Trp^7,Leu^8]-GnRH as compared with the [His^5,Trp^7,Tyr^8]-GnRH peptide. The two residues at which these forms differ are indicated in bold print. Degenerate 5' oligonucleotides used for PCR are shown below.

Fig. 2 shows the cDNA and predicted amino acid sequence of *H. burtoni* prepro[His^5,Trp^7,Tyr^8]-GnRH. The functional domains of prepro[His^5,Trp^7,Tyr^8]-GnRH are illustrated and compared to *H. burtoni* prepro[Trp^7,Leu^8]-GnRH. Amino acids are numbered beginning at the first residue in the GnRH decapeptide. The functional domains of both preprohormones are illustrated. Hydrophobic signal sequences (negative numbers) are directly followed by the GnRH decapeptide region (bold) and residues involved with posttranslational processing. The GAP peptide sequences follow the decapeptide and posttranslational processing sequences. A second processing site (underlined) indicates coding sequences for two novel peptides in the [His^5,Trp^7,Tyr^8]-GnRH preprohormone.

Fig. 3 shows the results of high stringency Northern blot analysis using total RNA from ventral, dorsal or whole brain, hybridized to ^{32}P-labelled prepro[His^5,Trp^7,Tyr^8]-GnRH riboprobe. Lanes 1, 2, and 3 were loaded with 5, 10 and 20 μg, respectively, of total RNA extracted from a pool of 3 brains. Lanes 4 through 7 contain 10 μg of total RNA each from one half of a single brain. Lanes 4 and 6 contain ventral brain RNA while lanes 5 and 7 contain RNA from the corresponding dorsal halves. One band of approximately 530 bases is seen in lanes containing ventral or total brain RNA. No signal was detected in total RNA extracted from dorsal brain halves, localizing the expression of this transcript exclusively to more ventral brain regions. Methylene blue staining of ribosomal bands confirmed that the equal amounts of total RNA were loaded into lanes 4-7 of the gel.

Fig. 4 shows localization of prepro [His^5,Trp^7,Tyr^8]-GnRH mRNA in *H. burtoni* brain. TOP: schematic view of a mid-sagittal section from *H. burtoni* brain showing the three populations of neurons immunoreactive for GnRH. For simplicity, the clusters are represented within the same plane although the mesencephalic population lies lateral to both the terminal nerve and POA populations. The terminal nerve GnRH population is indicated by is indicated by ▲, the POA population by ★, and the mesencephalic population by ●. The bottom six panels show the labelling with each of the two GnRH gene probes in the three brain regions with cells known to contain GnRH from immunocytochemistry. LEFT COLUMN: *in situ* hybridization with a prepro [His^5,Trp^7,Tyr^8]-GnRH digoxygenin-UTP riboprobe. Only the mesencephalic population was labeled. Scale bar = 100 μm. RIGHT COLUMN: *in situ* hybridization with a prepro [Trp^7,Leu^8]-GnRH digoxygenin-UTP riboprobe. Only the terminal nerve population was labeled. Scale bar = 100 μm. Note that neither probe labels the POA/hypothalamic group of cells.

Fig. 5 shows localization of prepro [Ser^8]-GnRH mRNA in *H. burtoni* brain. Midsagittal sections are shown in the panels. The top two panels show hybridization with a prepro [Ser^8]-GnRH riboprobe in subordinate and dominant male fish.

The bottom two panels show hybridization of the riboprobe in the terminal nerve area and the mesencephalon regions of the *H. burtoni* brain of dominant males.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

This invention provides isolated nucleic acid sequences encoding precursor proteins for the chicken II form of GnRH, which is referred to here as the [His^5,Trp^7,Tyr^8]-GnRH preprohormone, and for the [Ser^8]-GnRH preprohormone. The sequences can be used in a number of applications. For instance, they can be used for the recombinant production of the preprohormone polypeptides. In addition, transgenic animals (*e.g.*, fish and chickens, and non-human mammals) can be generated using the sequences provided here. Diagnostic assays for detecting nucleic acids encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone, as well as assays for these preprohormone proteins are also provided. These assays are particularly useful in the diagnosis of reproductive diseases or reproductive capacity in animals. Compositions and methods for using the sequences provided here are described in more detail below.

A) GnRH preprohormone polypeptides

As described above, there are eight different forms of GnRH that have been isolated and characterized from vertebrate species (see figure 1). A precursor for an additional form of GnRH, [Ser^8]-GnRH, is also described herein. Figure 1 shows all decapeptide sequences except [Ser^8]-GnRH with reference to the mammalian GnRH sequence. [Ser^8]-GnRH is a newly discovered form of GnRH, so that there are now 9 known forms of GnRH. Residues which differ from those in the mammalian form are identified both by position and residue. Thus, teleost GnRH is referred to as [Trp^7,Leu^8]-GnRH because, in teleosts, the Leu in position 7 is Trp and the Arg in position 8 is Leu when compared with the mammalian GnRH sequence. As can be seen in GnRH, positions 1,

2, 4, 9, and 10 are invariant, and positions 3 and 7 show only conservative changes.

The predicted amino acid sequence for the [His^5,Trp^7,Tyr^8]-GnRH preprohormone from the teleost fish, *H. burtoni*, is shown in Figure 2 (Seq. ID No. 2). In addition, the predicted amino acid sequence for the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone is shown in Seq ID No. 18. The precursor of a additional form of GnRH, [Ser^8]-GnRH, is shown in Seq. ID No. 20. The GnRH preprohormones have a characteristic structure consisting of three separate domains. The N-terminal domain consists of a signal peptide region which is a common feature of many secretory proteins. The second domain consists of the [His^5,Trp^7,Tyr^8]-GnRH decapeptide followed immediately by a 3 amino acid amidation and precursor processing site. The third, C-terminal, domain consists of a GnRH associated peptide or GAP peptide. The GAP peptide region of the [His^5,Trp^7,Tyr^8]-GnRH preprohormone may generate two peptides because of an additional potential proteolytic processing site in this region of the preprohormone. Thus, a single [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide consisting of the entire sequence of this region of the protein may be produced. Alternatively, there may be two [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptides present.

The [His^5,Trp^7,Tyr^8]-GnRH preprohormone and the [Ser^8]-GnRH preprohormone represent two members of a family of GnRH preprohormones. Homology between the *H. burtoni* and the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormones and the *H. burtoni* [Ser^8]-GnRH preprohormone was determined by the FASTA and GAP computer programs (version 7.3 Unix, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA). Using this procedure there is a 59% homology, and a 44% amino acid identity between the amino acid sequences of *H. burtoni* and treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormones. Also using this procedure, there is a 55% homology and a 33% identity between the amino acid sequences of the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone and the *H. burtoni* [Ser^8]-GnRH preprohormone.

The terms "[His^5,Trp^7,Tyr^8]-GnRH preprohormone", "[His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides" or "chicken II GnRH preprohormone" as used herein refer to all polypeptide precursor forms of [His^5,Trp^7,Tyr^8]-GnRH, but exclude the mature [His^5,Trp^7,Tyr^8]-GnRH decapeptide. Similarly, the terms "[Ser^8]-GnRH preprohormone" or "[Ser^8]-GnRH preprohormone polypeptides" as used herein refer to all polypeptide precursor forms of [Ser^8]-GnRH, but exclude the mature [Ser^8]-GnRH decapeptide. These terms also refer to biologically active fragments of the precursor proteins for [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH respectively. Significant biological activities include GnRH activity, latent GnRH activity and immunological activity. Latent GnRH activity means that the biologically active fragment can be proteolytically processed to yield active GnRH.

Immunological activity refers to immunoreactivity with an antibody raised to a full-length [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone or to segments of these preprohormones. A segment of a [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone will ordinarily comprise at least about 5 contiguous amino acids, typically at least about 7 contiguous amino acids, more typically at least about 9 contiguous amino acids, usually at least about 11 contiguous amino acids, preferably at least about 13 contiguous amino acids, more preferably at least about 16 contiguous amino acids, and most preferably at least about 20 to 30 or more contiguous amino acids from the preprohormone. Segments of a particular domain will be segments of the appropriate size within the corresponding domain. Biologically active fragments include the [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone GAP peptide or peptides.

The terms "GnRH related peptide", "GAP" or "GAP peptide" refer to the peptide of the C-terminal domain of a GnRH preprohormone, as described above for the [His^5,Trp^7,Tyr^8]-GnRH preprohormone and the [Ser^8]-GnRH preprohormone. Previously known GAP peptides are further described in Sherwood, N.A., *et al.*, *supra*. The terms

"[His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide" or "chicken II GnRH preprohormone GAP peptide" refer to the GAP peptide or peptides derived from the [His^5,Trp^7,Tyr^8]-GnRH preprohormone. The term "[Ser^8]-GnRH preprohormone GAP peptide" refers to the GAP peptide or peptides derived from the [Ser^8]-GnRH preprohormone. The [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone GAP peptides are produced by the proteolytic processing of [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone, respectively. The [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide is a biologically active [His^5,Trp^7,Tyr^8]-GnRH preprohormone fragment, as defined above, and is therefore included in the definition of [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides. Similarly, the [Ser^8]-GnRH preprohormone GAP peptide is a biologically active [Ser^8]-GnRH preprohormone fragment, as defined above, and is therefore included in the definition of [Ser^8]-GnRH preprohormone polypeptides.

The above-defined terms refer not only to the protein having the amino acid sequence disclosed here, but also to other proteins that are allelic, nonallelic or species variants of the *H. burtoni* and treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormones and the *H. Burtoni* [Ser^8]-GnRH preprohormone, as well as to natural or induced mutant forms of these proteins. For example, [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides generally show substantial sequence identity (determined as described above) to the amino acid sequence of the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in figure 2 or to the amino acid sequence of the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in Seq. ID No. 18. For example, fish [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides generally show substantial sequence identity to the amino acid sequence of the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in figure 2, and mammalian [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides generally show substantial sequence identity (determined as described above) to the amino acid sequence of the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in Seq. ID No. 18. Similarly, [Ser^8]-GnRH preprohormone polypeptides generally

show substantial sequence identity to the amino acid sequence of the *H. burtoni* [Ser^8]-GnRH preprohormone shown in Seq ID No. 20.

[His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides will also typically be specifically immunoreactive with antibodies raised against the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone of figure 2 or against the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in Seq. ID No. [???]. In particular, fish [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides will typically be specifically immunoreactive with antibodies raised against the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone of figure 2, and mammalian [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides will typically be specifically immunoreactive with antibodies raised against the treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone shown in Seq. ID No. 18. Similarly, [Ser^8]-GnRH preprohormone polypeptides will generally be specifically immunoreactive with antibodies raised against the *H. burtoni* [Ser^8]-GnRH preprohormone shown in Seq. ID No. 20.

A [His^5,Trp^7,Tyr^8]-GnRH preprohormone that specifically binds to or that is specifically immunoreactive to an antibody generated against a defined immunogen, such as an immunogen consisting of the amino acid sequence of Seq. ID No. 2 or Seq. ID No. 18 is determined in an immunoassay. The immunoassay uses a polyclonal antiserum which was raised to the protein of Seq. ID No. 2 or Seq. ID No. 18. This antiserum is selected to have low crossreactivity against GnRH preprohormones other than [His^5,Trp^7,Tyr^8]-GnRH preprohormones. Crossreactivity to these other forms of GnRH preprohormone is removed by immunoabsorption prior to use in the immunoassay.

In order to produce antisera to [His^5,Trp^7,Tyr^8]-GnRH precursor for use in an immunoassay, the proteins of Seq. ID No. 2 and Seq. ID No. 18 are isolated as described herein. For example, recombinant protein is produced in a mammalian cell line. An inbred strain of mice such as balb/c or rabbits are immunized with the protein of Seq. ID No. 2 or Seq. ID No. 18 using a standard adjuvant, such as Freund's adjuvant,

and a standard mouse immunization protocol (see Harlow and Lane, *supra*). Polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against GnRH preprohormones other than [His^5,Trp^7,Tyr^8]-GnRH preprohormone, using a competitive binding immunoassay such as the one described in Harlow and Lane, *supra*, at pages 570-573. Three non-[His^5,Trp^7,Tyr^8]-GnRH preprohormones are used in this determination: the mammalian GnRH preprohormone obtained from human, rat, and mouse (see Sherwood, *et al.* (1993), *supra*); the [Trp^7,Leu^8]-GnRH preprohormone obtained from Salmon and *H. burtoni*, and the [Ser^8]-GnRH preprohormone obtained from *H. burtoni*. These non-[His^5,Trp^7,Tyr^8]-GnRH preprohormones can be produced as recombinant proteins and isolated using standard molecular biology and protein chemistry techniques as described herein.

Immunoassays in the competitive binding format can be used for the crossreactivity determinations. For example, the protein of Seq. ID No. 2 or Seq. ID No. 18 can be immobilized to a solid support. Proteins added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to the protein of Seq. ID No. 2 or Seq. ID No. 18. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the proteins listed above are selected and pooled. Antisera raised against the protein of Seq ID No. 2 and the protein of Seq. ID No. 18 are pooled separately. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with the above-listed proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein to the immunogen protein. In order to make this comparison, the two proteins are each assayed at

a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. For example, if the amount of the second protein required is less than 10 times the amount of the protein of Seq. ID No. 2 required, then the second protein is said to specifically bind to an antibody generated to an immunogen consisting of the protein of Seq. ID No. 2.

A [Ser^8]-GnRH preprohormone that specifically binds to or that is specifically immunoreactive to an antibody generated against a defined immunogen, such as an immunogen consisting of the amino acid sequence of Seq. ID No. 20 can also determined in an immunoassay, by using a procedure similar to that described above for [His^5,Trp^7,Tyr^8]-GnRH preprohormones. In this case, however, polyclonal antisera are selected and tested for their cross reactivity against GnRH preprohormones other than [Ser^8]-GnRH preprohormone. The non-[Ser^8]-GnRH preprohormones used in this determination are: the mammalian GnRH preprohormone obtained from human, rat, and mouse (see Sherwood, *et al.* (1993), *supra*); the [Trp^7,Leu^8]-GnRH preprohormone obtained from Salmon and *H. burtoni*, and the [His^5,Trp^7,Tyr^8]-GnRH preprohormone obtained from *H. burtoni* and the treeshrew.

<u>B. Nucleic acid compositions for [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH preprohormone polypeptides</u>

This invention relates to isolated nucleic acid sequences encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone polypeptides. The nucleic acid compositions of this invention, whether RNA, cDNA, genomic DNA, or a hybrid of the various combinations, may be isolated from natural sources or may be synthesized *in vitro*. The nucleic acids claimed may be present in transformed or transfected whole cells, in a transformed or transfected cell lysate, or in a partially purified or substantially pure form.

The nucleic acid sequences of the invention are typically identical to or show substantial sequence identity (determined as described above) to the nucleic acid sequence

of SEQ ID. No. 1, Seq. ID No. 17, or Seq. ID No. 19. For example, nucleic acids encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides and which show substantial sequence identity to Seq. ID No. 1 will typically hybridize to the nucleic acid sequence of Seq. ID No. 1 under stringent conditions. Similarly, nucleic acids encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone polypeptides and which show substantial sequence identity to Seq. ID No. 17 will typically hybridize to the nucleic acid sequence of Seq. ID No. 17 under stringent conditions. With regard to [Ser^8]-GnRH preprohormones, nucleic acids encoding [Ser^8]-GnRH preprohormone polypeptides and which show substantial sequence identity to Seq. ID No. 19 will typically hybridize to the nucleic acid sequence of Seq. ID No. 19 under stringent conditions. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration is at least about 0.02 molar at pH 7 and the temperature is at least about 60°C. As other factors may significantly affect the stringency of hybridization, including, among others, base composition and size of the complementary strands, the presence of organic solvents and the extent of base mismatching, the combination of parameters is more important than the absolute measure of any one.

Techniques for nucleic acid manipulation of genes encoding these polypeptides such as subcloning nucleic acid sequences encoding polypeptides into expression vectors, labelling probes, DNA hybridization, and the like are described generally in Sambrook, *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989, which is incorporated herein by reference. This manual is hereinafter referred to as "Sambrook, *et al.*"

There are various methods of isolating the DNA sequences encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone polypeptides. For example, the DNA is isolated from a genomic or cDNA library using labelled oligonucleotide probes having sequences complementary to the sequences disclosed here. Restriction endonuclease digestion of genomic DNA or cDNA containing the [His^5,Trp^7,Tyr^8]-GnRH preprohormone gene or the [Ser^8]-GnRH preprohormone gene can be used to isolate nucleic acids encoding these proteins. Since the DNA sequences encoding [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH preprohormones are provided here, a panel of restriction endonucleases can be constructed to give cleavage of the DNA in the desired regions. After restriction endonuclease digestion, DNA encoding [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH preprohormones is identified by its ability to hybridize with nucleic acid probes, for example on Southern blots, and these DNA regions are isolated by standard methods familiar to those of skill in the art. *See* Sambrook, *et al.*

Various methods of amplifying target sequences, such as the polymerase chain reaction can also be used to prepare [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone DNA. Polymerase chain reaction technology (PCR) is used to amplify nucleic acid sequences of the [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone polypeptides directly from mRNA, from cDNA, and from genomic libraries or cDNA libraries. The isolated sequences encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone may also be used as templates for PCR amplification.

Appropriate primers and probes for amplifying nucleic acids encoding nucleic acids encoding the [His^5,Trp^7,Tyr^8]-GnRH or the [Ser^8]-GnRH preprohormone polypeptides can be generated from analysis of the DNA sequences. In brief, oligonucleotide primers complementary to the two 3' borders of the DNA region to be amplified are synthesized. The polymerase chain reaction is then carried out using the two primers. *See PCR Protocols: A Guide to Methods and Applications.* (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), Academic Press, San Diego (1990).

Primers can be selected to amplify the entire regions encoding the full-length [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormones or to amplify smaller DNA segments of these preprohormones as desired.

Oligonucleotides for use as probes are chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage, S.L. and Carruthers, M.H., 1981, *Tetrahedron Letts.*, 22(20):1859-1862 using an automated synthesizer, as described in Needham-VanDevanter, D.R., *et al.*, 1984, *Nucleic Acids Res.*, 12:6159-6168. Purification of oligonucleotide is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson, J.D. and Regnier, F.E., 1983, *J. Chrom.*, 255:137-149.

The sequence of the synthetic oligonucleotide can be verified using the chemical degradation method of Maxam, A.M. and Gilbert, 1980, in W., Grossman, L. and Moldave, D., eds. Academic Press, New York, *Methods in Enzymology*, 65:499-560.

Other methods known to those of skill in the art may also be used to produce and isolate nucleic acids encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone. *See* Sambrook, *et al.* for a description of other techniques for the isolation of DNA encoding specific protein molecules.

C. Expression of [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone.

Once the DNA encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone is isolated and cloned, one may express these preprohormones in a variety of recombinantly engineered cells. It is expected that those of skill in the art are knowledgeable in the numerous expression systems available for expression of the DNA encoding these preprohormones. No attempt to describe in detail the various methods known for the expression of proteins in prokaryotes or eukaryotes is made here.

In brief summary, the expression of natural or synthetic nucleic acids encoding [His^5,Trp^7,Tyr^8]-GnRH or

[Ser^8]-GnRH preprohormone will typically be achieved by operably linking the DNA or cDNA to a promoter (which is either constitutive or inducible), followed by incorporation into an expression vector. The vectors can be suitable for replication and integration in either prokaryotes or eukaryotes. Typical expression vectors contain transcription and translation terminators, initiation sequences, and promoters useful for regulation of the expression of the polynucleotide sequence encoding [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone polypeptides. To obtain high level expression of a cloned gene, such as those polynucleotide sequences encoding [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone, it is desirable to construct expression plasmids which contain, at the minimum, a strong promoter to direct transcription, a ribosome binding site for translational initiation, and a transcription/translation terminator. The expression vectors may also comprise generic expression cassettes containing at least one independent terminator sequence, sequences permitting replication of the plasmid in both eukaryotes and prokaryotes, *i.e.*, shuttle vectors, and selection markers for both prokaryotic and eukaryotic systems. *See* Sambrook *et al.* Examples of expression of [His^5,Trp^7,Tyr^8]-GnRH and [Ser^8]-GnRH preprohormone polypeptides in both prokaryotic and eukaryotic systems are described below.

1. Expression in Prokaryotes

A variety of procaryotic expression systems may be used to express [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides. Examples include *E. coli*, *Bacillus*, *Streptomyces*, and the like. For example, [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides may be expressed in *E. coli*. As another example, these preprohormone polypeptides may be expressed in fish cells. Other fish proteins such as rainbow trout growth hormone have been successfully expressed in *E. coli* (*See, e.g.*, Agellon *et al.* (1986) *DNA* 5:463-477 and U.S. Patent No, 4,849,359).

It is essential to construct expression plasmids which contain, at the minimum, a strong promoter to direct transcription, a ribosome binding site for translational initiation, and a transcription/translation terminator. Examples of regulatory regions suitable for this purpose in *E. coli* are the promoter and operator region of the *E. coli* tryptophan biosynthetic pathway as described by Yanofsky, C., 1984, *J. Bacteriol.*, 158:1018-1024 and the leftward promoter of phage lambda (P_L) as described by Herskowitz, I. and Hagen, D., 1980, *Ann. Rev. Genet.*, 14:399-445. The inclusion of selection markers in DNA vectors transformed in *E. coli* is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline, or chloramphenicol. *See* Sambrook *et al.* for details concerning selection markers for use in *E. coli*.

[His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH preprohormone polypeptides produced by prokaryotic cells may not necessarily fold properly. During purification from *E. coli*, the expressed protein may first be denatured and then renatured. This can be accomplished by solubilizing the bacterially produced proteins in a chaotropic agent such as guanidine HCl and reducing all the cysteine residues with a reducing agent such as beta-mercaptoethanol. The protein is then renatured, either by slow dialysis or by gel filtration. *See* U.S. Patent No.4,511,503.

Detection of the expressed antigen is achieved by methods known in the art as radioimmunoassays, or Western blotting techniques or immunoprecipitation. Purification from *E. coli* can be achieved following procedures described in U.S. Patent No. 4,511,503.

2. Expression in Eukaryotes

A variety of eukaryotic expression systems such as yeast, insect cell lines, bird, fish, and mammalian cells, are known to those of skill in the art. As explained briefly below, [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides may also be expressed in these eukaryotic systems.

Synthesis of heterologous proteins in yeast is well known and described. *Methods in Yeast Genetics*, Sherman, F., *et al.*, Cold Spring Harbor Laboratory, (1982) is a well recognized work describing the various methods available to produce the preprohormone in yeast.

Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or other glycolytic enzymes, and an origin of replication, termination sequences and the like as desired. For instance, suitable vectors are described in the literature (Botstein, *et al.*, 1979, *Gene*, 8:17-24; Broach, *et al.*, 1979, *Gene*, 8:121-133).

Two procedures are used in transforming yeast cells. In one case, yeast cells are first converted into protoplasts using zymolyase, lyticase or glusulase, followed by addition of DNA and polyethylene glycol (PEG). The PEG-treated protoplasts are then regenerated in a 3% agar medium under selective conditions. Details of this procedure are given in the papers by J.D. Beggs, 1978, *Nature* (London), 275:104-109; and Hinnen, A., *et al.*, 1978, *Proc. Natl. Acad. Sci. USA*, 75:1929-1933. The second procedure does not involve removal of the cell wall. Instead the cells are treated with lithium chloride or acetate and PEG and put on selective plates (Ito, H., *et al.*, 1983, *J. Bact.*, 153:163-168).

[His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides, once expressed, can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The monitoring of the purification process can be accomplished by using Western blot techniques or radioimmunoassays of other standard immunoassay techniques.

The sequences encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides can also be ligated to various expression vectors for use in transforming cell cultures of, for instance, mammalian, insect, bird or fish origin. Illustrative of cell cultures useful for the production of the polypeptides are mammalian cells. Mammalian cell systems often will be in the form of

monolayers of cells although mammalian cell suspensions may also be used. A number of suitable host cell lines capable of secreting intact proteins have been developed in the art, and include the CHO cell lines, various human cells such as COS cell lines, HeLa cells, myeloma cell lines, Jurkat cells, etc. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter (*e.g.*, a HSV *tk* promoter or *pgk* (phosphoglycerate kinase) promoter), an enhancer (Queen *et al.* (1986) *Immunol. Rev.* 89:49), and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites (*e.g.*, an SV40 large T Ag poly A addition site), and transcriptional terminator sequences.

Other animal cells useful for production of the preprohormone are available, for instance, from the American Type Culture Collection Catalogue of Cell Lines and Hybridomas (7th edition, 1992). For example, fish cells that can be used include cells derived from rainbow trout, salmon, and the like. Suitable promoters and vectors are described for instance, in Friedenreich *et al.*, (1990) *Nuc. Acids Res.* 18:3299-3305.

Appropriate vectors for expressing preprohormone polypeptides in insect cells are usually derived from the SF9 baculovirus. Suitable insect cell lines include mosquito larvae, silkworm, armyworm, moth and *Drosophila* cell lines such as a Schneider cell line (*See* Schneider *J. Embryol. Exp. Morphol.* 27:353-365 (1987).

As indicated above, the vector, *e.g.*, a plasmid, which is used to transform the host cell, preferably contains DNA sequences to initiate transcription and sequences to control the translation of the preprohormone gene sequence. These sequences are referred to as expression control sequences.

As with yeast, when higher animal host cells are employed, polyadenlyation or transcription terminator sequences from known mammalian genes need to be incorporated into the vector. An example of a terminator sequence is the polyadenlyation sequence from the bovine growth hormone gene.

Sequences for accurate splicing of the transcript may also be included. An example of a splicing sequence is the VP1 intron from SV40 (Sprague, J. *et al.*, 1983, J. Virol. 45: 773-781).

Additionally, gene sequences to control replication in the host cell may be incorporated into the vector such as those found in bovine papilloma virus type-vectors. Saveria-Campo, M., 1985, "Bovine Papilloma virus DNA a Eukaryotic Cloning Vector" in *DNA Cloning Vol. II a Practical Approach* Ed. D.M. Glover, IRL Press, Arlington, Virginia pp. 213-238.

The host cells are competent or rendered competent for transformation by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation, fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, DEAE dextran, electroporation and micro-injection of the DNA directly into the cells.

The transformed cells are cultured by means well known in the art. *Biochemical Methods in Cell Culture and Virology*, Kuchler, R.J., Dowden, Hutchinson and Ross, Inc., (1977). The expressed preprohormone polypeptides are isolated from cells grown as suspensions or as monolayers. The latter are recovered by well known mechanical, chemical or enzymatic means.

D. <u>Purification of [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides</u>

The polypeptides produced by recombinant DNA technology may be purified by standard techniques well known to those of skill in the art. Recombinantly produced preprohormone polypeptides can be directly expressed or expressed as a fusion protein. The protein is then purified by a combination of cell lysis (*e.g.*, sonication) and affinity chromatography. For fusion products, subsequent digestion of the fusion protein with an appropriate proteolytic enzyme release the desired polypeptide.

The polypeptides of this invention may be purified to substantial purity by standard techniques well known in the art, including selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. *See*, for instance, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982), incorporated herein by reference.

E. Production of [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides by protein chemistry techniques

The polypeptides of the invention can be synthetically prepared in a wide variety of ways. For instance polypeptides of relatively short size, can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. *See*, for example, Stewart and Young, *Solid Phase Peptide Synthesis*, 2d. ed., Pierce Chemical Co. (1984).

F. Modification of nucleic acid and polypeptide sequences

The nucleotide sequences used to transfect the host cells used for production of recombinant preprohormone polypeptides can be modified according to standard techniques to yield preprohormone polypeptides with a variety of desired properties. The polypeptides of the present invention can be readily designed and manufactured utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the polypeptides can vary from the naturally-occurring sequence at the primary structure level by amino acid insertions, substitutions, deletions, and the like. These modifications can be used in a number of combinations to produce the final modified protein chain.

The amino acid sequence variants can be prepared with various objectives in mind, including facilitating purification and preparation of the recombinant polypeptides. The modified polypeptides are also useful in, for example, modifying plasma half-life, improving therapeutic efficacy,

and lessening the severity or occurrence of side effects during therapeutic use. The amino acid sequence variants are usually predetermined variants not found in nature but exhibit the same immunogenic activity as naturally occurring polypeptides. For instance, polypeptide fragments comprising only a portion (usually at least about 60-80%, typically 90-95%) of the primary structure may be produced. In general, modifications of the sequences encoding the polypeptides may be readily accomplished by a variety of well-known techniques, such as site-directed mutagenesis (*See*, Gillman and Smith, *Gene* 8:81-97 (1979) and Roberts, S. *et al.*, *Nature* 328:731-734 (1987)).

G. Production of Transgenic Non-human Animals

The invention also encompasses methods and polynucleotide constructs which are employed for generating transgenic non-human animals which express the preprohormone polypeptides of the invention. The constructs are used to produce transgenic non-human mammals, birds (*e.g.*, chickens), fish and the like. Incorporation of nucleotide sequences of the invention allow modification of the reproductive and other characteristics of the transgenic animals. For example, fish that do not readily spawn in captivity can be induced to reproduce.

Appropriate constructs and methods for production of transgenic animals are known. Typically, the coding sequence of interest is operably linked to expression regulatory sequences. In such transgenes, the expression regulatory sequence is at least the minimal sequences required for efficient cell-type specific expression, which generally are at least a promoter and sequences upstream of the promoter, which provide for efficient expression in the target cells. Usually the sequences upstream of the promoter are used contiguously, although various deletions and rearrangements can be employed. Some desired regulatory elements (*e.g.*, enhancers, silencers) may be relatively position-insensitive, so that the regulatory element will function correctly even if positioned differently in a transgene than in the

corresponding germline gene. For example, an enhancer may be located at a different distance from a promoter, in a different orientation, and/or in a different linear order. For example, an enhancer that is located 3' to a promoter in germline configuration might be located 5' to the promoter in a transgene.

Typically, expression regulation sequences are chosen to produce tissue-specific or cell type-specific expression of the recombinant DNA. Once a tissue or cell type is chosen for expression, expression regulation sequences are chosen. Generally, such expression regulation sequences are derived from genes that are expressed primarily in the tissue or cell type chosen. Preferably, the genes from which these expression regulation sequences are obtained are expressed substantially only in the tissue or cell type chosen, although secondary expression in other tissue and/or cell types is acceptable if expression of the recombinant DNA in the transgene in such tissue or cell type is not detrimental to the transgenic animal.

Particularly preferred expression regulation sequences are those endogenous to the species of animal to be manipulated. However, expression regulation sequences from other species such as those from human genes may also be used. In some instances, the expression regulation sequences and the recombinant DNA sequences (either genomic or cDNA) are from the same species. Alternatively, the expression regulation sequences and recombinant DNA sequences (either cDNA or genomic) are obtained from different species. In such cases, the expression regulation and recombinant DNA sequence are heterologous to each other.

In certain embodiments, it is desirable to use gene targeting, mediated by homologous recombination between a targeting polynucleotide construct and a homologous chromosomal sequence, to replace an endogenous gene with the gene encoding a mutant of the preprohormone gene. Methods and materials for preparing such constructs are known by those of skill in the art and are described in various references. *See*, e.g., Thomas et al., *Cell* 51:503 (1987) and Capecchi,

Science 244:1288 (1989). Homologous targeting constructs have at least one region having a sequence that substantially corresponds to, or is substantially complementary to, a predetermined endogenous target gene sequence (e.g., an exon sequence, an enhancer, a promoter, an intronic sequence, or a flanking sequence of the target gene). Such a homology region serves as a template for homologous pairing and recombination with substantially identical endogenous gene sequence(s). In the targeting of transgenes, such homology regions typically flank the replacement region, which is a region of the targeting transgene that is to undergo replacement with the targeted endogenous gene sequence. Thus, a segment of the targeting transgene flanked by homology regions can replace a segment of the endogenous gene sequence by double crossover homologous recombination.

The constructs described above are introduced into pluripotent or totipotent cells using standard techniques. Briefly, this technology involves the insertion of the desired transgene construct into an appropriate cell line (*e.g.*, mammalian embryonic stem (ES) cells or fertilized oocytes) that is capable of differentiating into germ cell tissue. Methods of introducing transgenes into embryonal target cells include microinjection of the transgene into the pronuclei of fertilized oocytes or nuclei of ES cells of the non-human animal. Such methods for murine species are well known to those skilled in the art. Alternatively, the transgene may be introduced into an animal by infection of zygotes with a retrovirus containing the transgene (Jaenisch, R. (1976) *Proc. Natl. Acad. Sci. USA* 73:1260-1264).

The production of transgenic non-human mammals (*e.g.*, mice, cows, pigs, and the like) are described in International Application No. WO/08216 and Krimpenfort, *et al.*, *Biotechnology* 9:844-847 (1991). Germ line transgenesis of both chickens and japanese quail has been described. For a general description of the production of transgenic birds, *See*, Shuman (1991) *Experientia* 47:897-905.

In certain preferred embodiments, transgenic fish are generated. A variety of examples of the production of

transgenic fish have been reported (*See, e.g.*, Guyomard *et al.* (1989) *Biochimie* 71:857-863 and Rokkones *et al.* (1989) *J. Comp. Physiol. B.* 158:751-785. Vectors can be constructed using mammalian promoters, *e.g.*, from the murine metallothionen gene, or viral promoters, *e.g.*, Rous sarcoma virus and simian virus 40 (*see, e.g.*, Stuart *et al.* (1988) *Development* 103:403-412 and Yoon *et al.* (1990) *Aquaculture* 85:21-33). Alternatively, vectors comprising entirely piscine regulatory sequences can be used (Liu *et al.* (1990) *Bio/Technology* 8:1268-1272).

The appropriate construct comprising an expression cassette containing the sequences encoding the preprohormone is introduced by microinjection into recently fertilized eggs isolated from spawning fish of the appropriate species (*e.g.*, salmon, trout, and the like). The injected embryos are reared in the appropriate medium and allowed to develop into mature fish using standard aquacultural techniques.

Tissue samples from the transgenic animals can be analyzed using standard techniques for detecting the presence or absence of a target sequence. For instance, Fluorescent In Situ Hybridization (FISH) can be used to detect the transgene. Several guides to FISH techniques are available, *e.g.*, Gall *et al. Meth. Enzymol.*, 21:470-480 (1981) and Angerer *et al.* in *Genetic Engineering: Principles and Methods* Setlow and Hollaender, Eds. Vol 7, pgs 43-65 (Plenum Press, New York 1985).

The sequences can also be detected by PCR using primers and probes specific for the transgene. Standard PCR methods useful in the present invention are described in *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, eds., Academic Press, San Diego).

H. Pharmaceutical Compositions

The preprohormone polypeptides (*e.g.*, GAP peptides) of the present invention can also be administered to control reproductive function. Thus, the polypeptides can be prepared in pharmaceutical compositions and administered using methods well known in the art. For instance, the pharmaceutical

compositions can be administered orally with feed for the regulation of reproductive function in farm animals such as chickens. The compositions can also be suitable for parenteral administration. The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include powder, tablets, pills, and capsules.

Suitable pharmaceutical formulations for use in the present invention are found in *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985). A variety of pharmaceutical compositions comprising preprohormone polypeptides of the present invention and pharmaceutically effective carriers can be prepared.

Pharmaceutical compositions of the invention include nucleic acid sequences encoding the preprohormone inserted into a suitable gene therapy vector. A variety of different gene therapy vectors may be used. Current strategies and vectors for gene therapy are reviewed in Miller A.D, (1992) *Nature*, 357:455-460 and Mulligan, R.C. (1993) *Science* 260:926-932. The gene therapy vector may be delivered to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion). For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous injection. Suitable formulations for injection are found in *Remington's Pharmaceutical Sciences*, *supra*. The pharmaceutical compositions are suitable in a variety of drug delivery systems. For a brief review of present methods of drug delivery, *See*, Langer, *Science* 249:1527-1533 (1990) which is incorporated herein by reference.

<u>I. Detection of Nucleic acids encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides and detection of [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone polypeptides by immunoassays</u>

The present invention provides methods for detecting DNA or RNA encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone and for measuring these preprohormones by immunoassay techniques. These methods are useful for two general purposes. First, assays for detection of nucleic acids encoding these preprohormones are important for the isolation these nucleic acids from a variety of species. As described above, it is known that the mature [His^5,Trp^7,Tyr^8]-GnRH hormone is present in a wide range of species (*See* Sherwood, N.M. *et al., supra*). Now that the [His^5,Trp^7,Tyr^8]-GnRH preprohormone and [Ser^8]-GnRH cDNA's have been cloned and isolated, nucleic acids encoding these preprohormones may be isolated from a variety of species according to the methods described in section (B) above and by use of the nucleic acid hybridization assays described below. The immunoassays described below may be useful for isolation of nucleic acids encoding the [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormones by expression cloning methods. See section (B) above and Sambrook, *et al.*

The assays described below are also useful as *in vitro* diagnostic assays to determine the reproductive status of farm animals or of fish in aquaculture. For example, testing of fish for [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone production at various times may play a role in the regulation of fish reproduction in aquaculture. In addition, the discovery of the [His^5,Trp^7,Tyr^8]-GnRH mature hormone in placental mammals (*See* Sherwood, N.M., *supra*) and the demonstration, herein, of the [His^5,Trp^7,Tyr^8]-GnRH preprohormone in the treeshrew suggests that the [His^5,Trp^7,Tyr^8]-GnRH preprohormone may also be produced in humans. Therefore, the assays described below may also be useful in tests for reproductive function in humans.

1. Nucleic Acid Hybridization Assays

A variety of methods for specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art. *See* Sambrook, *et al.* For example, one method for evaluating the presence or absence of [His^5,Trp^7,Tyr^8]-GnRH preprohormone DNA or [Ser^8]-GnRH preprohormone DNA in a sample involves a Southern transfer. Briefly, the digested genomic DNA is run on agarose slab gels in buffer and transferred to membranes. Hybridization is carried out using the nucleic acid probes discussed above. As described above, nucleic acid probes are designed based on the known nucleic acid sequences encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone. Visualization of the hybridized portions allows the qualitative determination of the presence or absence of DNA encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone.

Similarly, a Northern transfer may be used for the detection of mRNA encoding [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone. In brief, the mRNA is isolated from a given cell sample using an acid guanidinium-phenol-chloroform extraction method. The mRNA is then electrophoresed to separate the mRNA species and the mRNA is transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled probes are used to identify the presence or absence of the [His^5,Trp^7,Tyr^8]-GnRH or the [Ser^8]-GnRH preprohormone transcript.

A variety of nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in "*Nucleic Acid Hybridization, A Practical Approach*," Ed. Hames, B.D. and Higgins, S.J., IRL Press, 1985; Gall and Pardue (1969), *Proc. Natl. Acad. Sci.*, U.S.A., 63:378-383; and John, Burnsteil and Jones (1969) *Nature*, 223:582-587.

For example, sandwich assays are commercially useful hybridization assays for detecting or isolating nucleic acid

sequences. Such assays utilize a "capture" nucleic acid covalently immobilized to a solid support and labelled "signal" nucleic acid in solution. The clinical sample will provide the target nucleic acid. The "capture" nucleic acid and "signal" nucleic acid probe hybridize with the target nucleic acid to form a "sandwich" hybridization complex. To be effective, the signal nucleic acid cannot hybridize with the capture nucleic acid.

Typically labelled signal nucleic acids are used to detect hybridization. Complementary nucleic acids or signal nucleic acids may be labelled by any one of several methods typically used to detect the presence of hybridized polynucleotides. The most common method of detection is the use of autoradiography with ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P-labelled probes or the like. Other labels include ligands which bind to labelled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labelled ligand.

Detection of a hybridization complex may require the binding of a signal generating complex to a duplex of target and probe polynucleotides or nucleic acids. Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal.

The label may also allow indirect detection of the hybridization complex. For example, where the label is a hapten or antigen, the sample can be detected by using antibodies. In these systems, a signal is generated by attaching fluorescent or enzyme molecules to the antibodies or in some cases, by attachment to a radioactive label. (Tijssen, P., "Practice and Theory of Enzyme Immunoassays," *Laboratory Techniques in Biochemistry and Molecular Biology*, Burdon, R.H., van Knippenberg, P.H., Eds., Elsevier (1985), pp. 9-20.)

The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system which multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction

(PCR) system and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBA™, Cangene, Mississauga, Ontario) and Q Beta Replicase systems.

An alternative means for determining the level of expression of the [His^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone gene or the [Ser^{8}]-GnRH preprohormone is *in situ* hybridization. *In situ* hybridization assays are well known and are generally described in Angerer, *et al.*, *Methods Enzymol.*, 152:649-660 (1987). In an *in situ* hybridization assay, cells, preferentially bovine lymphocytes are fixed to a solid support, typically a glass slide. If DNA is to be probed, the cells are denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of labeled probes specific to [His^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone or the [Ser^{8}]-GnRH preprohormone. The probes are preferably labelled with radioisotopes or fluorescent reporters.

2. Production of Antibodies and Development of Immunoassays

In addition to detecting expression of the [His^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone or the [Ser^{8}]-GnRH preprohormone by nucleic acid hybridization, one can also use immunoassays to detect the preprohormones. Immunoassays can be used to qualitatively or quantitatively analyze for the preprohormone or to specifically detect the GAP peptide. A general overview of the applicable technology can be found in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Pubs., N.Y. (1988), incorporated herein by reference.

a. Antibody Production

A number of immunogens may be used to produce antibodies specifically reactive with the [His^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone or the [Ser^{8}]-GnRH preprohormone. Recombinant preprohormone is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Naturally occurring preprohormone may also be used either in pure or impure form. Synthetic peptides made using the

[His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone sequences described herein may also used as an immunogen for the production of antibodies to the preprohormone. Preferentially, recombinant [His^5,Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone or fragments thereof are expressed in bacterial cells as described above, and purified as generally described above. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the preprohormone.

Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the [His^5,Trp^7,Tyr^8]-GnRH preprohormone or the [Ser^8]-GnRH preprohormone. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera is prepared. Further fractionation of the antisera to enrich for antibodies reactive to the preprohormone can be done if desired. (*See* Harlow and Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*See*, Kohler and Milstein, *Eur. J. Immunol.* 6:511-519 (1976), incorporated herein by reference). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Methods of production of synthetic peptides are known to those of skill ı the art. Briefly, the predicted immunogenic regions of s^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone or [Ser^{8}]-GnRH preprohormone sequences described herein are identified. For production of antibodies specific to the [His^{5},Trp^{7},Tyr^{8}]-GnRH or [Ser^{8}]-GnRH preprohormone GAP peptides, predicted imm nogenic regions of the preprohormone sequence in the GAP region of the molecule are used (see Seq. ID No. 2 , Seq ID No. 18 and Seq. ID No 20. Peptides preferably at least 10 amino acids in length are synthesized corresponding to these regions and the peptides are conjugated to larger protein molecules for subsequent immunization. Production of monoclonal or polyclonal antibodies is then carried out as described above.

b. <u>Immunoassays</u>

Antibodies reactive with a particular protein can be measured by a variety of immunoassay methods. For a review of immunological and immunoassay procedures in general, *See Basic and Clinical Immunology* 7th Edition (D. Stites and A. Terr ed.) 1991. Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay*, E.T. Maggio, ed., CRC Press, Boca Raton, Florida (1980); "Practice and Theory of Enzyme Immunoassays," P. Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers B.V. Amsterdam (1985); and, Harlow and Lane, *Antibodies, A Laboratory Manual*, *supra*, each of which is incorporated herein by reference.

Immunoassays for measurement of [His^{5},Trp^{7},Tyr^{8}]-GnRH preprohormone or [Ser^{8}]-GnRH preprohormone polypeptides can be performed by a variety to methods known to those skilled in the art. In brief, immunoassays to measure the preprohormones or the GAP peptides can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample analyte competes with a labeled analyte for specific binding sites on a capture agent bound to a solid surface. Preferably the capture agent is an antibody specifically reactive with the [His^{5},Trp^{7},Tyr^{8}]-GnRH

preprohormone or $[Ser^8]$-GnRH preprohormone produced as described above. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

In a competitive binding immunoassay, $[His^5,Trp^7,Tyr^8]$-GnRH preprohormone or $[Ser^8]$-GnRH preprohormone present in the sample competes with labelled preprohormone for binding to a specific binding agent, for example, an antibody specifically reactive with the preprohormone. The binding agent may be bound to a solid surface to effect separation of bound labelled preprohormone from the unbound labelled preprohormone. Alternately the competitive binding assay may be conducted in liquid phase and any of a variety of techniques known in the art may be used to separate the bound labelled preprohormone from the unbound labelled preprohormone. Following separation, the amount of bound labeled preprohormone is determined. The amount of preprohormone present in the sample is inversely proportional to the amount of labelled preprohormone binding.

Alternatively, a homogenous immunoassay may be performed in which a separation step is not needed. In these immunoassays, the label on the preprohormone is altered by the binding of the preprohormone to its specific binding agent. This alteration in the labelled protein results in a decrease or increase in the signal emitted by label, so that measurement of the label at the end of the immunoassay allows for detection or quantitation of the preprohormone.

$[His^5,Trp^7,Tyr^8]$-GnRH preprohormone or $[Ser^8]$-GnRH preprohormone may also be determined by a variety of noncompetitive immunoassay methods. For example, a two-site, solid phase sandwich immunoassay is used. In this type of assay, a binding agent for the preprohormone, for example an antibody, is attached to a solid phase. A second preprohormone binding agent, which may also be an antibody, and which binds the preprohormone at a different site, is labelled. After binding at both sites on the preprohormone has occurred, the unbound labelled binding agent is removed and the amount of labelled binding agent bound to the solid

phase is measured. The amount of labelled binding agent bound is directly proportional to the amount of preprohormone in the sample.

Western blot analysis can also be done to determine the presence of [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-GnRH preprohormone in a sample. Electrophoresis is carried out, for example, on a tissue sample suspected of containing the preprohormone. Following electrophoresis to separate the proteins, and transfer of the proteins to a suitable solid support such as a nitrocellulose filter, the solid support is then incubated with an antibody reactive with the preprohormone. This antibody may be labeled, or alternatively may be it may be detected by subsequent incubation with a second labelled antibody that binds the anti-preprohormone antibody.

The immunoassay formats described above employ labelled assay components. The label can be in a variety of forms. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide variety of labels may be used. The component may be labeled by any one of several methods. Traditionally a radioactive label incorporating ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P was used. Non-radioactive labels include ligands which bind to labelled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labelled ligand. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation. For a review of various labelling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

This invention also embraces diagnostic kits for detecting the presence of [His^5, Trp^7,Tyr^8]-GnRH preprohormone or [Ser^8]-GnRH preprohormone in tissue or blood samples which comprise a container containing antibodies selectively immunoreactive to the preprohormone and instructional material for performing the test. This invention further embraces diagnostic kits for detecting [His^5,Trp^7,Tyr^8]-GnRH or [Ser^8]-

GnRH preprohormone DNA or RNA in tissue or blood samples which comprise nucleic probes as described herein and instructional material.

EXAMPLES

Example 1: Isolation of a cDNA encoding *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone

A) Development of an oligonucleotide probe specific for *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone cDNA

To obtain nucleotide sequences encoding a [His^5,Trp^7,Tyr^8]-GnRH-like peptide from *H. burtoni*, we employed a novel polymerase chain reaction (PCR) strategy using brain regions known to contain GnRH neuronal populations (see Fig. 4 for a schematic view of *H. burtoni* brain). Animals were sacrificed and whole brains removed as described in White, S.A. & Fernald, R.D. *J. Neurosci.* (1993) 13:434-441. The dorsal portion of the telencephalon, the optic tectum and cerebellum from each brain was removed. Total RNA was isolated from the remaining ventral portions by guanidine thiocyanate-acid phenol extraction and converted into cDNA using reverse transcriptase (Superscript, BRL; Gaithesburg, MD). cDNA synthesis was primed with 0.5 μg of a bipartite oligonucleotide consisting of a homopolymer of 9 d(T) residues at the 3' end. The 5' domain comprised a sequence of 14 nucleotides which included a restriction endonuclease recognition site. The sequence of the oligonucleotide is: 5' GCAGAAGCTTCAGCT(9) 3' (Seq. ID No. 15).

The resulting cDNA was used as substrate for nested PCR. In both rounds of amplification, the downstream primer was equivalent to the 5' domain of the bipartite oligonucleotide, described above. In the first PCR reaction, the upstream primer was a pool of 24 14-mers representing all possible coding sequences for the first five amino acids of [His^5,Trp^7,Tyr^8]-GnRH (see figure 1(b)). 12.5 pmol of the upstream primer was used. After 40 cycles of 94°C/30 s; 55°C/30s; 72°C/15s, 0.05 μl of the product of this reaction was aliquoted to a second reaction using 12.5 pmol of a new upstream primer together with the same downstream primer under the same cycling conditions. The new primer pool of 24 16-mers consisted of all possible coding combinations for residues 5-9 of the decapeptide (see figure 1(b)). Reactions

were performed using Taq polymerase (synthesized courtesy of Dr. R. Moses) on a Perkin-Elmer 9600 thermocycler. Oligonucleotides were synthesized as needed using an Applied Biosystems 391 instrument.

PCR products were separated using a 1.5% GTG agarose gel (FMC, Rockville, ME, USA) and products of greater than 300 bases were electrocuted and subcloned into M13 for sequencing as described in Bond, C.T., Francis, R.C., Fernald, R.D. & Adelman, J.P. (1991) *J. Mol. Endocrinol.* 5:931-937. Nucleotide sequences were analyzed for the presence of the second upstream primer (ending in the codon for amino acid 9; see Fig. 2), followed by a codon for the final conserved amino acid of the decapeptide, ^{10}Gly, and sequences for the canonical GnRH amidation and peptide processing site, ^{11}Gly^{12}Lys^{13}Arg. (*See* Douglass, J., Civelli, O. & Herbert, E. (1984) *Ann. Rev. Biochem.* 53:665-715.)

A radiolabeled nucleotide probe 29 bases in length was derived from a putative [His^5,Trp^7,Tyr^8]-GnRH PCR product. The sequence of this oligonucleotide is GGGAATGCAGCTACCTGAGACCCCAGAGG (Seq. ID No. 16).

B) Isolation of a full-length cDNA encoding the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH preprohormone

To obtain full-length coding sequences, poly-adenylated RNA was isolated from *H. burtoni* ventral brain regions (see above) and used to construct a cDNA library in λgt10. 250,000 primary recombinant were screened (Genescreen; NEN-Dupont, Boston, MA, USA) using the radiolabeled oligonucleotide of Seq ID No. 16, described above. Positively hybridizing phage were purified by successive rescreens at reduced density. The cDNA inserts were subcloned into M13 phage and the nucleotide sequences determined. A cDNA containing the full-length coding sequence was subcloned into a transcription vector (Pselect, USB) for generation of sense and antisense riboprobes.

Example 2: Sequence Analysis of the *H. burtoni* cDNA encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone

Six positively hybridizing clones were identified from the *H. burtoni* ventral brain cDNA library as described above and the nucleotide sequences of the inserts were determined. The full-length cDNA sequence is shown in Seq. ID No. 1, and the predicted amino acid sequence for the *H. burtoni* chicken II GnRH preprohormone is shown in figure 2 (Seq. ID No. 2). The longest open reading frame containing the [His^5,Trp^7,Tyr^8]-GnRH coding sequence begins with a methionine initiator codon and predicts a 90 amino acid protein (Fig. 2). The first 23 residues are largely hydrophobic and are likely a signal peptide as is found in many polyprotein neuroendocrine precursors (see Douglass, J., Civelli, O. & Herbert, E. (1984) *Ann. Rev. Biochem.* 53:665-715.) The signal sequence is in direct linkage with [His^5,Trp^7,Tyr^8]-GnRH, which is followed by $Gly^{11}Lys^{12}Arg^{13}$. These residues follow the decapeptide sequence in all of the other cloned GnRH preprohormones and serve as substrates for post-translational processing (*See* Douglass, J., *et al.*, *supra*). The amino acid sequence of the [His^5,Trp^7,Tyr^8]-GnRH preprohormone (Seq. ID No. 3) is shown for comparison in Figure 2. The remainder of the precursor may code for two additional peptides also generated by proteolytic processing, neither of which show any homology to sequences in protein databases. One peptide, 28 residues in length, is followed by dibasic residues (arg-arg, see Fig. 2 legend), and the remaining 19 amino acids could comprise a second processing product. There is only about 10% identity between the two *H. burtoni* GnRH preprohormones when the GnRH motif is excluded (see Figure 2).

Example 3: Northern blot analysis of *H. burtoni* brain tissue for the [His^5,Trp^7,Tyr^8]-GnRH Precursor

Northern blots were prepared as previously described (see Bond, C.T., Francis, R.C., Fernald, R.D. & Adelman, J.P. (1991) *J. Mol. Endocrinol.* 5:931-937) and probed with antisense riboprobes synthesized with SP6 RNA polymerase,

incorporating ^{32}P-labelled UTP to a specific activity of 3 x 10^8 dpm/μg. Hybridizations contained approximately 10^7 dpm per ml. High stringency conditions were obtained using 50% formamide at 65°C with washes in 0.1 x SSC at 70°C.

Northern blot analysis of total RNA extracted from dorsal, ventral and whole *H. burtoni* brain revealed a single transcript of *ca.* 530 bases expressed in ventral and whole brain but absent from dorsal structures (Fig. 3). Radiolabelled antisense riboprobes for both the [His^5,Trp^7,Tyr^8]-GnRH and [Trp^7,Leu^8]-GnRH preprohormones were prepared and used to probe Northern blots of polyA+ RNA isolated from total brain at both high and low stringencies. The results (not shown) indicate that the probes for the two preprohormones do not cross-hybridize at either stringency with the alternative transcript.

Example 4: *In situ* hybridization

The site of [His^5,Trp^7,Tyr^8]-GnRH mRNA production in *H. burtoni* brain was assessed using *in situ* hybridization. Animals were sacrificed, brains removed and tissue fixed as previously described (*See* White, *et al.*, *J. Neurosci.* (1993) 13:434-44). 40 μm cryostat sections were mounted on poly L-lysine subbed slides for hybridization to the digoxygenin UTP (DIG, Boehringer Mannheim) -labeled [His^5,Trp^7,Tyr^8]-GnRH riboprobe. Hybridizations were carried out at high stringency (55°C in 50% formamide) according to the manufacturer's instructions with the following modifications: 150 μl of hybridization solution was applied to slides containing 4 to 5 brain sections. Following hybridization, 3 washes were done in 2 X SSC at 55°C for one hour each, followed by one wash in 1 X SSC for 30 min at room temperature. To prepare the riboprobe, subcloned prepro[His^5,Trp^7,Tyr^8]-GnRH was linearized with Eco RI (GIBCO-BRL, Grand Island, New York, USA), followed by filling in of recessed 3' termini with the Klenow fragment of DNA polymerase I and Proteinase K treatment. The template DNA was purified using Magic DNA Clean-Up System (Promega Corporation, Madison, Wisconsin, USA) and 300 *ng* was used for transcription of a riboprobe with SP6

RNA polymerase (GIBCO-BRL) in the presence of DIG-conjugated UTP. Final concentrations in a 100μl reaction volume were: 1 X SP6 reaction buffer, 1 X NTP labeling mixture (Boehringer-Mannheim, Indianapolis, Indiana, USA), 4.0 mM DTT, 80 units RNAsin (Promega), 13.2 μg BSA, and 120 units of SP6 polymerase. Solutions for both Northern analysis and *in situ* hybridizations were prepared using water treated with 0.075% diethylpyrocarbonate.

[His^5,Trp^7,Tyr^8]-GnRH mRNA expression, localized using *in situ* hybridization, was found only in a cluster on neurons found in the mesencephalon (Fig. 4 (TOP)) previously shown to contain GnRH using immunocytochemistry. Davis, M.R. & Fernald, R.D., *J. Neurobiol.* (1990) 21:1180-1188. Interestingly, [His^5,Trp^7,Tyr^8]-GnRH is not expressed in the hypothalamic nucleus (Fig. 4 (LEFT COLUMN)) which projects to the pituitary and hence probably does not directly influence gonadotropin activation. The probes for the two species of GnRH found in *H. burtoni* do not cross-react either when applied *in situ* (see Fig. 4) or in Northern blots (see above). Thus distinct forms of GnRH encoding genes are expressed in distinct brain regions.

In *H. burtoni*, as in other vertebrates, three neuronal populations have been shown to contain GnRH using immunocytochemistry: the terminal nerve, the hypothalamic/preoptic area and the mesencephalon. Davis, M.R. & Fernald, R.D., *J. Neurobiol.* (1990) 21:1180-1188; White, S.A. & Fernald, R.D., *J. Neurosci.* (1993) 13:434-441; see Fig. 4, TOP). Here we have shown that [His^5,Trp^7,Tyr^8]-GnRH mRNA is expressed only in the mesencephalic population (Fig. 4, (LEFT COLUMN)). In contrast, [Trp^7,Leu^8]-GnRH mRNA is localized only in the terminal nerve nucleus of the telencephalon (Fig. 4 (RIGHT COLUMN)). Furthermore, the lack of *in situ* hybridization within the GnRH-immunoreactive cells of the hypothalamic/preoptic area (Fig. 4) reveals that the GnRH-encoding gene expressed there is sufficiently different to elude detection by either probe. Thus, despite its potent releasing factor activity (*cf.* Ngamvongchon, S. Rivier, J.E. & Sherwood, N.M. (1992), *Regul. Pept.* 42:63-73)),

[His^5,Trp^7,Tyr^8]-GnRH is unlikely to influence pituitary function directly in *H. burtoni*. Moreover, this suggests that yet a third gene encoding for GnRH may be responsible for regulating reproduction directly.

The structure of the [His^5,Trp^7,Tyr^8]-GnRH predicts two novel peptides in addition to it, in contrast to the single associated peptide predicted by [Trp^7,Leu^8]-GnRH mRNA (see Figure 2). The lack of homology between the two known GnRH preprohormones in *H. burtoni* suggests that these genes have either evolved independently or diverged from a common ancestor well before the appearance of teleost fish.

The novel preprohormone structure and extrahypothalamic location of the *H. burtoni* [His^5,Trp^7,Tyr^8]-GnRH precursor may indicate a unique function for this GnRH form and/or its associated peptides. While hypothalamic GnRH clearly functions as a releasing hormone, neuromodulatory and neurotransmitter roles have been postulated for GnRHs in other brain regions (reviewed in Loumaye, E., Thorner, J. & Catt, K.J. (1982) *Science* 218:1323-1325). The localization of the [His^5,Trp^7,Tyr^8]-GnRH mRNA to the mesencephalon supports previous immunological studies in which [His^5,Trp^7,Tyr^8]-GnRH-like peptides were found concentrated in caudal brain areas. Such mesencephalic GnRH-containing neurons may play a role in the central organization of reproductive behavior as suggested for poecilid fish where a similar neuronal population has been described. *See* Miller, K.E. & Kriebel, R.M. (1986) *Gen. Comp. Endocrinol.* 64:396-400. These cells contain GnRH and project to a spinal neurosecretory group which, in turn, has been postulated to play a role in gonadal duct contractility.

As noted, only one GnRH peptide (previously called mammalian GnRH), that which regulates the pituitary gonadotropes, has been demonstrated in more recently evolved mammals. This form can be detected in some bony fish, but is absent in *H. burtoni* as well as in reptiles and birds (reviewed in Loumaye, E., Thorner, J. & Catt, K.J. (1982) *Science* 218:1323-1325, *supra*). In contrast, [His^5,Trp^7,Tyr^8]-GnRH has widespread expression among vertebrates with the notable exception of more derived mammals. Since non-

mammalian animals express more than one GnRH peptide, it is possible that many mammals may express more than one form of GnRH, particularly the [Ser^8]-GnRH form.

Example 5: Isolation of a cDNA encoding a treeshrew [His^5,Trp^7,Tyr^8]-GnRH preprohormone

To isolate the sequence encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone isoform in the treeshrew *Tupaia bulangeri*, we used a polymerase chain reaction (PCR) strategy similar to that used in the isolation of this isoform in a teleost fish, *Haplochromis burtoni*, as described Example 1, herein. The amino acid sequence of the [^{5}His,^{7}Trp, ^{8}Try]GnRH decapeptide guided selection of primers for PCR amplification. Treeshrew brain mRNA was primed with a bipartite oligonucleotide for cDNA synthesis. The 3' domain of this oligonucleotide consisted of a homopolymer of d(T)24 residue, while the 5' domain comprised a known sequence of 10 nucleotides which included a restriction endonuclease recognition site. The resulting cDNA was used as substrate for nested PCR. The downstream primer in both rounds of amplification was bipartite. A randomly generated hexamer including all possible combinations of nucleotides at seven positions composed the 3' end and ten known nucleotides, including an Hind III restriction endonuclease recognition site, made up the 5' end. In the first PCR reaction, the upstream primer was a pool of 24 tetradecamers representing all possible coding sequences for the first five amino acids of [His^5,Trp^7,Tyr^8]-GnRH. The sequence of these primers is represented in the following two sequences: GCACGAATTCCA A/G CA T/C TGGTCNCA (Seq. ID No. 21) and GCACGAATTCCA A/G CA T/C TGGAG T/C CA (Seq. ID No. 22). The product of the first PCR reaction served as substrate for nested reactions where the new upstream primer was bipartite. The 5' end of this primer contained a known restriction enzyme recognition sequence while the 3' end comprised a pool of 24 tetradecamers representing all possible coding combinations for residues 5-8 of the decapeptide. These sequence this pool of primers is: GCACGAATTCCA T/C GGNTGGTA T/C CC (Seq. ID No. 23).

PCR products were separated by electrophoresis in a 2.5% GTG agarose gel (FMC) and products > 250 bases electrocuted and subcloned into M13 phage vector for sequence analysis (Bond *et al.*, 1991). Nucleotide sequences were analyzed for the presence of the second upstream primer followed by the codons for the final amino acid of the decapeptide, Gly, and the next three amino acids Gly, Lys, and Arg. These residues follow the decapeptide sequence in the other cloned GnRH preprohormones and serve as substrates for posttranslational processing. One clone contained these landmarks followed by an open reading frame of 165 (or 219) nucleotides 55 (or 73) amino acids. The numbers in parentheses refer to the nucleotide and amino acid length if the alternate codon using methionine is used as the start codon. To confirm the identity of this clone as the [His^5,Trp^7,Tyr^8]-GnRH preprohormone, a downstream region of the clone was chosen from which to generate new primers for 5' rapid amplification of cDNA ends (RACE). See Frohman *et al.* in (1993) *Methods in Enzymology* Volume 218, Wu, R., editor, Academic Press, Inc., San Diego, California, USA for a more detailed description of RACE methodology.

Substrate cDNA for the 5' RACE was synthesized from mRNA primed with randomly generated hexamers including all possible combinations of nucleotides at 7 positions. These cDNA products were tailed at their 5' ends with dATP in a terminal deoxytransferase reaction (GIBCO BRL, Grand Island, New York, USA). Two rounds of PCR amplification were performed on the tailed cDNA in which the 3' antisense primer consisted of a bipartite oligonucleotide (poly d(T)17 plus Hind III recognition site) designed to anneal to the 5' terminal poly d(A)s. The outer 5' antisense primer for the first round of PCR consisted of a septadecamer corresponding to a region 111-127 nucleotides downstream of the decapeptide region. The oligonucleotide sequence for the outer primer was CAGGGCACACTGTCCTC (Seq. ID No. 24). For the second round of amplification, the inner 5' antisense primer was bipartite, composed of a known sequence of ten nucleotides including a restriction endonuclease cleavage site plus another 17

nucleotides 72-88 nucleotides downstream of the decapeptide region of the preprohormone candidate. The oligonucleotide sequence for the inner primer was GCACGAATTCCGAAGCTATGAGCAGTC (Seq. ID No. 25). PCR products were prepared for sequencing (as above) and screened for the presence of the innermost primer followed by antisense sequences from the candidate clone and by antisense codons for the GnRH decapeptide. A majority of the generated clones were positive for the [His^5,Trp^7,Tyr^8]-GnRH encoding region and extended approximately 135 more nucleotides before ending in the poly d(T)/Hind III primer-sequence. Only one clone was isolated containing preprohormone sequence 3' to that region amplified by 5' RACE.

Example 6: Sequence Analysis of the treeshrew cDNA encoding the [His^5,Trp^7,Tyr^8]-GnRH preprohormone

Only one clone was isolated in Example 5 which contained preprohormone sequence 3' to that region amplified by 5' RACE. To confirm the sequence, we designed a nondegenerate bipartite oligonucleotide whose 5' end consisted of ten nucleotides including a restriction endonuclease cleavage site and whose 3' end encoded residues 3-8 of the decapeptide region. The sequence of the oligonucleotide was GCACGAATTCGGTCCCACGGCTGGTAC (Seq. ID No. 26). Poly d(T) primed cDNA was used as substrate in PCR amplification between the above oligonucleotide and a bipartite poly d(T)/Hind III primer downstream. Products were prepared, sequenced and screened as before resulting in many positive clones containing the same sequence as that found in the primary downstream clone and ending in polyadenylated tails.

The fullest sequence of this preprohormone was obtained by aligning overlapping regions from at least ten consensus 5' RACE templates with at least ten consensus sequences generated in the first and second 3' PCR amplifications. Sequences were aligned (Color Alignment Macros, M. Haygood) and differences attributed to PCR-error were assessed. The deduced amino acid sequence is shown in Seq. ID No. 18. As with GnRH preprohormones previously

isolated, a signal sequence precedes the decapeptide, followed by a conserved proteolytic processing site and an associated peptide. (*See Comparative Biology and Evolutionary Relationships of Treeshrews*, W.P. Luckett, ed., Plenum Press, New York (1980).) Treeshrews are part of the super order *Archonta*, which includes primates, flying lemurs, bats and elephant shrews. Treeshrews are thus very closely related to primates, yet primitive. As a consequence, they are very frequently used to discover information about molecules, processes, etc., which may link much more primitive rodents with primates. Thus, [Ser^8]-GnRH preprohormone may be present in primates including humans.

Example 7: Isolation of a cDNA encoding *H. burtoni* [Ser^8]-GnRH preprohormone

To obtain nucleotide sequences encoding a [Ser^8]-GnRH-like peptide from *H. burtoni*, a novel polymerase chain reaction (PCR) strategy using brain regions known to contain GnRH neuronal populations was employed (see Fig. 4 for a schematic view of *H. burtoni* brain). Animals were sacrificed and whole brains removed as described in White, S.A. & Fernald, R.D. *J. Neurosci.* (1993) 13:434-441. *H. burtoni* mRNA was isolated from the preoptic area, brain tissue known to contain the GnRH neuronal population which projects to the pituitary. cDNA synthesis was primed with a bipartite oligonucleotide whose 3' domain consisted of a homopolymer of d(T)9 residues, while the 5' domain comprised a known sequence of 14 nucleotides which included a restriction endonuclease recognition site. The resulting cDNA was used as substrate for nested PCR. In both rounds of amplification, the downstream primer was as described above. In the first PCR reaction, the 5' end of the primer contained a known restriction enzyme recognition sequence and the 3' sequence was a mixture of 48 tetradecamers representing all possible coding sequences for the first five amino acids of [Ser^8]GnRH. (See Powell, J.F.F., *et al.* Regulatory Peptides, submitted.) The sequence of the this pool of oligonucleotides was as follows: GCACGAATTCCA A/G CA T/C TGGTCNTA T/C GG (Seq. ID No.

27). The product of this reaction served as substrate for nested reactions where the new upstream primer was bipartite. The 5' end of this primer also contained a known restriction enzyme recognition sequence while the 3' end comprised a pool of 384 tetradecamers representing all possible coding combinations for residues 5-8 of the decapeptide. The sequences of these oligonucleotides are represented in the following two sequences: GCACGAATTCTA T/C GGN C/T TNTCNCC (Seq. ID No. 28) and GCACGAATTCTA T/C GGN C/T TNAG C/T CC (Seq. ID No. 29).

PCR products were separated by electrophoresis in a 2% GTG agarose gel (FMC Bioproducts, Rockville, Maine, USA) and products >250 bases were electroeluted and subcloned into M13 phage vector for sequence analysis (see Bond *et al.* (1991) *J. Mol. Endocrinol.*, 5:931-937). Nucleotide sequences were analyzed for the presence of the second upstream primer followed by the codons for the final amino acid of the decapeptide, glycine, and the following three amino acid residues gly, lys, arg (see figure 2). These residues follow the decapeptide sequence in the other cloned GnRH preprohormones and serve as substrates for posttranslational processing. Several independent clones from 3 separate first-round and 7 second-round PCR reactions contained these landmarks followed by an open reading frame of 201 nucleotides (67 amino acids). To confirm the identity of these clones as the [Ser^8]-GnRH preprohormone, templates were visually compared after alignment in a spreadsheet (Color Alignment Macros, M. Haygood, Scripps Institute of Oceanography, San Diego, California, USA) and a region of consensus was chosen from which to generate new primers for 5' rapid amplification of cDNA ends (RACE). See Frohman *et al.* in (1993) *Methods in Enzymology* Volume 218, Wu, R., editor, Academic Press, Inc., San Diego, California, USA for a more detailed description of RACE methodology.

The bipartite primer used to make the substrate cDNA for the 5' RACE reactions consisted of a randomly generated hexamer including all possible combinations of nucleotides at 7 positions. These cDNA products were tailed at their 5' ends

with dATP in a terminal deoxytransferase reaction (GIBCO BRL, Grand Island, New York, USA). Two rounds of PCR amplification were performed on the tailed cDNA in which the 3' antisense primer consisted of a bipartite oligonucleotide (poly d(T)17 plus Bam HI recognition site), designed to anneal to the 5' terminal poly d(A)s. The outer 5' antisense primer for the first round of PCR consisted of 18 nucleotides corresponding to a region of high consensus amongst the candidate preprohormone templates. The sequence of this primer is CATATTGCCCAGTGTGTC (Seq. ID No. 30). For the second round of amplification, the inner 5' antisense primer was bipartite, composed of a known sequence of 14 nucleotides including a restriction endonuclease cleavage site plus another 18 nucleotides from the preprohormone candidate which were 3' to but did not overlap with the first selected sequence. The nucleotide sequence of this primer is GCACGAATTCGTGTCTGAGAAGTTGTCC (Seq. ID No. 31). PCR products were prepared for sequencing as described above, and were screened for the presence of the innermost primer followed by antisense sequences from the candidate templates and by antisense codons for the GnRH decapeptide itself.

Example 8: Sequence Analysis of the cDNA encoding *H. burtoni* [Ser^8]-GnRH preprohormone

Approximately 80% of the generated clones were positive for the [Ser^8]-GnRH-encoding region and extended approximately 140 more nucleotides before ending in the poly d(T)/BamHI primer sequence. The fullest sequence of this preprohormone was obtained by aligning overlapping regions from consensus 5' RACE templates with consensus sequences generated in the primary PCR amplification. Sequences were aligned (Color Alignment Macros) and differences attributed to PCR-error were assessed. The sequence of the full length cDNA is shown as Seq. ID No. 19 and the deduced amino acid sequence is shown as Seq. ID. No. 20. In each case, a signal sequence precedes one of the decapeptide forms, followed by a conserved proteolytic processing site and an associated peptide.

Example 9: Localization of [Ser^8]-GnRH preprohormone mRNA expression in *H. Burtoni* brain sections by *in situ* hybridization

The identity of the [Ser^8]-GnRH preprohormone obtained above as that encoding the GnRH contained within preoptic area GnRH-ir neurons which govern gonadotropin secretion from the pituitary was confirmed through *in situ* analysis on *H. burtoni* brain sections. These neurons are known to enlarge or shrink in size depending on the social and reproductive state of the fish (see Davis & Fernald (1991) *J. Neurobiol.* 21:1180-1188). Tissue sections were prepared from both a dominant and a subordinate animal, which contain either large or small GnRH immunoreactive (GnRH-ir) neurons, respectively. Social status was determined by behavioral observations. Animals were sacrificed, brains removed and tissue fixed as previously described in Example 4, herein. Cryostat sections (40μm) were mounted on poly(L-lysine)-coated slides for hybridization.

A hybridization probe was generated by PCR amplification of a consensus 5' RACE product cloned into M13 using lac and rev lac primers. Following gel electrophoresis performed as described in Example 7, herein, a single band of the correct size (ca. 210 bases) was electroeluted and fragments were subcloned into pSK+ (Stratagene, San Diego, California, USA). Double-stranded sequencing confirmed that several clones contained the 5' RACE insert and one was chosen as a template for generation of an antisense digoxygenin-UTP (Boehringer Mannheim, Indianapolis, Indiana, USA) riboprobe. The riboprobe was prepared as follows: The plasmid was linearized with BamHI (GIBCO/BRL, Grand Island, New York, USA) and purified using the Wizard clean-up system from Promega, Madison, Wisconsin, USA). 300 ng was used for transcription of an RNA probe with T7 polymerase (GIBCO/BRL, Grand Island, New York, USA) and a 1xNTP labelling mixture (Boehringer-Mannheim, Indianapolis, Indiana, USA) as described in Example 4, herein.

Hybridizations were carried out at high stringency (60°C in 60% formamide) under conditions slightly modified

from those suggested by the digoxygenin manufacturer (Boehringer Mannheim, Indianapolis, Indiana, USA). 150 μl of hybridization solution was applied to slides containing five to six brain sections. After hybridization, 2 washes were done in 2x SSC, followed by washes in 1.5x and 1x SSC respectively at 60°C with a final wash at room temperature in 1x SSC. All washes were for 20 minutes. Slides were dehydrated, coverslipped and viewed using a light microscope.

Figure 5 shows a panel of midsagittal sections from *H. burtoni* brain, focusing on the three regions which contain GnRH-ir cell populations (see Davis & Fernald (1990), supra). The top three regions were hybridized to the [Ser^8]-GnRH riboprobe as described above. Within this top tier, the first section is split, showing preoptic area cells labeled with the [Ser^8]-GnRH riboprobe from both a dominant and a subordinate male fish. They reveal that the [Ser^8]-GnRH transcript is expressed in the preoptic region within a cell population that exhibits neuronal size plasticity correlated with social state. The bottom two panels show hybridization of the riboprobe in the terminal nerve area and the mesencephalon regions of the *H. burtoni* brain of dominant males. Thus, in *H. burtoni*, the three different genes encoding GnRH are expressed in three distinct regions of the brain.

While not wishing to be bound by theory, the *in situ* hybridization data supports an important reproductive role for the hypothalamic-preoptic form of GnRH, [Ser^8]-GnRH, in the cichlid fish *H. burtoni*. In this species, preoptic GnRH-ir cells grow and shrink with social and reproductive state. Dominant males have large GnRH-producing cells, large testes, and are reproductively active. Subordinate males have small GnRH-producing cells, small testes and do not mate. Preoptic GnRH gene expression may therefore not only reflect the reproductive state of the animal, but may also be responsive to social cues.

Also while not wishing to be bound by theory, the distribution of [His^5,Trp^7,Tyr^8]-GnRH mRNA in *H. Burtoni* brain, coupled with other information, suggests that this form of GnRH may coordinate reproductive behavior with reproductive

state. [His^5,Trp^7,Tyr^8]-GnRH is the most widespread and ancient GnRH isoform, having been immunologically detected in caudal brain regions of all non-mammalian vertebrates studied to date, often localized to a midbrain neuronal population. While in goldfish these cells project to the median eminence and result in the delivery of [His^5,Trp^7,Tyr^8]GnRH (in addition to the hypothalamic form) to the pituitary, this isoform is not present in the pituitary of *H. burtoni*. Neither is it widely detected in the pituitaries of other species. This fact stands in contrast to the marked efficacy of the [His^5,Trp^7,Tyr^8]GnRH isoform in causing release of gonadotropins in cultures of dispersed pituitary cells. Thus, while this isoform is capable of acting at pituitary GnRH receptors, it seldom has the *in vivo* opportunity to demonstrate its functional potency. However, because of this biological activity, nucleic acids encoding the [His^5,Trp^7,Tyr^8]GnRH precursor, as well as the nucleic acids encoding the [Ser^8]-GnRH precursor, may be useful in the production of transgenic animals, including transgenic fish.

<u>Example 9: Northern blot analysis of *H. burtoni* brain tissue for the [Ser^8]-GnRH Precursor</u>

To determine the size of the [Ser^8]-GnRH mRNA precursor relative to the sizes of the other GnRH preprohormones in *H. burtoni*, Northern analyses were performed at high stringency as described in Example 3, herein. Poly(A)+ RNA was isolated from ventral and dorsal brain regions using the FastTrak kit from Invitrogen Corporation, San Diego, California. 1.5 μg aliquots were loaded onto replicate lanes of a denaturing 2.5% agarose gel for size separation. The gel was prepared as a Northern blot using capillary transfer. Individual lanes were subjected to high-stringency hybridizations (50% formamide, 62°C) with riboprobes for each of the 3 GnRH preprohormones in *H. burtoni*, either alone and in combination. Riboprobes were synthesized with T7 polymerase (GIBCO BRL, Grand Island, New York, USA), incorporating [a-32P]UTP (Amersham Corporation, Arlington Heights, Illinois, USA) to a specific activity of 3

X 108 dpm/μg. Hybridization mixtures contained 107 dpm per ml. The [Ser^8]-GnRH riboprobe labeled a band of approximately 550 bases while the size of the [Trp^7,Leu^8]GnRH and [His5,Trp7,Tyr8]GnRH transcripts are approximately 500 and 590 bases, respectively.

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: The Board of Trustees of Leland Stanford, Jr. University,
And
State of Oregon, acting by and through The Oregon State Board of Higher Education on behalf of The Oregon Health Sciences University

(B) STREET: 900 Welch Road, Suite 350
(C) CITY: Palo Alto
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) POSTAL CODE (ZIP): 94304-1858
(G) TELEPHONE: (415) 723-0651
(H) TELEFAX: (415) 725-7295
(I) TELEX:

(ii) TITLE OF INVENTION: Nucleic acids Encoding [His-5,Try-7,Tyr-8]-GnRH Preprohormone and [Ser-8]-GnRH preprohormone and Their Uses

(iii) NUMBER OF SEQUENCES: 31

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: WO not yet designated
(B) FILING DATE:
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/147,771
(B) FILING DATE: 05-NOV-1993

(vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Weber, Kenneth A.
(B) REGISTRATION NUMBER: 31,677
(C) REFERENCE/DOCKET NUMBER: 14210-000400PC

(viii) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 543-9600
(B) TELEFAX: (415) 543-5043

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 563 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Haplochromis burtoni

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 135..389

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 243..389
(D) OTHER INFORMATION: /note= "Nucleotide sequence that encodes the GAP peptide."

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..563
(D) OTHER INFORMATION: /note= "cDNA sequence that contains nucleotide sequence that encodes GnRH{5-His,7-Trp,8-Tyr} from Haplochromis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GACTGTCAGC GCAACTGGAT TTTAGCACTA AATCCACCAA AGGAAAAGAA CATTTTGAAG      60

TGAACCCCTG CAGGTACACT GAGGGAAACT TGGACTGATA AAGCTGTGAA ATCTAAGACT     120

AAGGTGGGAA TATC ATG TGT GTG TCT CGA CTG GCT TTG CTC TTG GGG CTG       170
                Met Cys Val Ser Arg Leu Ala Leu Leu Leu Gly Leu
                 1               5                  10

CTT CTC TGT GTG GGG GCT CAG CTG TCC TTT GCC CAG CAC TGG TCC CAT       218
Leu Leu Cys Val Gly Ala Gln Leu Ser Phe Ala Gln His Trp Ser His
        15                  20                  25

GGT TGG TAT CCT GGA GGA AAA AGG GAG CTG GAC TCC TTT GGC ACA TCA       266
Gly Trp Tyr Pro Gly Gly Lys Arg Glu Leu Asp Ser Phe Gly Thr Ser
    30                  35                  40

GAG ATT TCA GAG GAG ATT AAG CTG TGT GAA GCA GGG GAA TGC AGC TAC       314
Glu Ile Ser Glu Glu Ile Lys Leu Cys Glu Ala Gly Glu Cys Ser Tyr
 45                  50                  55                  60

CTG AGA CCC CAG AGG AGG AGT ATC CTG AGA AAC ATT CTT CTG GAT GCC       362
Leu Arg Pro Gln Arg Arg Ser Ile Leu Arg Asn Ile Leu Leu Asp Ala
                65                  70                  75

TTA GCC AGA GAG CTT CAG AAG AGA AAG TGACATCTTT CCAGAGCCTC             409
Leu Ala Arg Glu Leu Gln Lys Arg Lys
            80                  85

TTTTCTATAG TAACCCACTT CCCTTTGTAT TTCTGCCTTG ACGTGATTTT GTGATCATCT     469

GGCCTTGCTG TTTGTAATGT TTGTCAGTAA ATTTGTCCTG TTTTTTTCGA TGTGAAAATT     529

GTGTCCCAAA ATAAATATCT ATTTTTATAT TAAA                                 563
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO. 2:

```
Met Cys Val Ser Arg Leu Ala Leu Leu Leu Gly Leu Leu Leu Cys Val
 1              5                   10                  15

Gly Ala Gln Leu Ser Phe Ala Gln His Trp Ser His Gly Trp Tyr Pro
            20                  25                  30

Gly Gly Lys Arg Glu Leu Asp Ser Phe Gly Thr Ser Glu Ile Ser Glu
        35                  40                  45

Glu Ile Lys Leu Cys Glu Ala Gly Glu Cys Ser Tyr Leu Arg Pro Gln
    50                  55                  60

Arg Arg Ser Ile Leu Arg Asn Ile Leu Leu Asp Ala Leu Ala Arg Glu
65                  70                  75                  80

Leu Gln Lys Arg Lys
                85
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Haplochromis burtoni

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..90
(D) OTHER INFORMATION: /note= "Name: PreproGnRH{7-Trp,8-Leu}."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Glu Ala Gly Ser Arg Val Ile Met Gln Val Leu Leu Leu Ala Leu
1               5                   10                  15

Val Val Gln Val Thr Leu Ser Gln His Trp Ser Tyr Gly Trp Leu Pro
            20                  25                  30

Gly Gly Lys Arg Ser Val Gly Glu Leu Glu Ala Thr Ile Arg Met Met
        35                  40                  45

Gly Thr Gly Gly Val Val Ser Leu Pro Asp Glu Ala Asn Ala Gln Ile
    50                  55                  60

Gln Glu Arg Leu Arg Pro Tyr Asn Ile Ile Asn Asp Asp Ser Ser His
65                  70                  75                  80

Phe Asp Arg Lys Lys Arg Phe Pro Asn Asn
                85                  90
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mammal

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Glu His Trp Ser Tyr Gly Leu Arg Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Chicken (I)

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH{8-Gln}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Glu His Trp Ser Tyr Gly Leu Gln Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Salmon

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH{7-Trp,8-Leu}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Glu His Trp Ser Tyr Gly Trp Leu Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Catfish

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH{5-His,8-Asn}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Glu His Trp Ser His Gly Leu Asn Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Chicken (II)

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH{5-His,7-Trp,8-Tyr}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Glu His Trp Ser His Gly Trp Tyr Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Dogfish

(ix) FEATURE:
(A) NAME/KEY: Peptide

(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note= "GnRH{5-His,7-Trp,8-Leu}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Glu His Trp Ser His Gly Trp Leu Pro Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Lamprey

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note=
"GnRH{3-Tyr,5-Leu,6-Glu,7-Trp,8-Lys}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Glu His Tyr Ser Leu Glu Trp Lys Pro Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Lamprey

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /note=
"GnRH{3-Tyr,5-His,6-Asp,7-Trp,8-Lys}"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Glu His Tyr Ser His Asp Trp Lys Pro Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Alpha-yeast mating factor

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His Trp Leu Glu Leu Lys Pro Gly
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CARCAYTGGT CNCA 14

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CAYGGNTGGT AYCC 14

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCAGAAGCTT CAGCT 15

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
GGGAATGCAG CTACCTGAGA CCCCAGAGG                                29
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 422 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 72..338
(D) OTHER INFORMATION: /product= "The [His-5,Trp-7,Tyr-8]-GnRH preprohormone from Treeshrew"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 60..62
(D) OTHER INFORMATION: /note= "Encodes potential second initiator methionine."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
CTCGAGCCAG AAGGAGCCAC CCTACCTATA GCCTCTTCCA CTGCTGTCCC CGTCCTGCCA      60

TGGCCAGTTC C ATG CTG GGC TTC CTC CTC CTG CTG CTG CTG CTG ATG GCT       110
             Met Leu Gly Phe Leu Leu Leu Leu Leu Leu Leu Met Ala
               1               5                  10

GCC CAC CCT GGA CCC TCG GAG GCC CAG CAT TGG TCC CAC GGC TGG TAC        158
Ala His Pro Gly Pro Ser Glu Ala Gln His Trp Ser His Gly Trp Tyr
     15                 20                  25

CCT GGA GGA AAG CGA GCC TCC AAC TCA CCC CAG GAC CCT CAA AGT GCC        206
Pro Gly Gly Lys Arg Ala Ser Asn Ser Pro Gln Asp Pro Gln Ser Ala
 30                 35                  40                  45

CTT AGG CCC CCA GCC CCC AGC GCA GCC AGA CTG CTC ATA GCT TCC GAA        254
Leu Arg Pro Pro Ala Pro Ser Ala Ala Arg Leu Leu Ile Ala Ser Glu
                 50                  55                  60

GCG CTG CTC TGG CTT CCC CCG AGG ACA GTG TGC CCT GGG AGG GCA GGA        302
Ala Leu Leu Trp Leu Pro Pro Arg Thr Val Cys Pro Gly Arg Ala Gly
             65                  70                  75
```

```
CCA CAG CAG GAT GGT CTC TCC GCA GGA AGC AGC ACC TGATGCGGAC          348
Pro Gln Gln Asp Gly Leu Ser Ala Gly Ser Ser Thr
        80                      85

ACTGCTGAGC GCAGCCGGAG CGCCGCGCCC CGCCGCCGT[illegible] [illegible]CAATAAAGC CGTGAGATTC   408

CCGAAAAAAA AAAA                                                     422
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met Leu Gly Phe Leu Leu Leu Leu Leu Leu Leu Met Ala Ala His Pro
  1               5                  10                  15

Gly Pro Ser Glu Ala Gln His Trp Ser His Gly Trp Tyr Pro Gly Gly
             20                  25                  30

Lys Arg Ala Ser Asn Ser Pro Gln Asp Pro Gln Ser Ala Leu Arg Pro
         35                  40                  45

Pro Ala Pro Ser Ala Ala Arg Leu Leu Ile Ala Ser Glu Ala Leu Leu
     50                  55                  60

Trp Leu Pro Pro Arg Thr Val Cys Pro Gly Arg Ala Gly Pro Gln Gln
 65                  70                  75                  80

Asp Gly Leu Ser Ala Gly Ser Ser Thr
                 85
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 491 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 71..364
(D) OTHER INFORMATION: /product= "The [Ser-8]-GnRH preprohormone from H. burtoni"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 47..49
(D) OTHER INFORMATION: /note= "Encodes potential second initiator methionine."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
TTAGACTTCA CAAAGGACAG CAGAAGAGAT CAGAAGTTCT TGTCTAATGC ACAGAAGCTT    60
```

```
TATCCTCAGA ATG GCT GCA AAA ATC TTG GCA CTG TGG CTG CTG CTC GCA      109
           Met Ala Ala Lys Ile Leu Ala Leu Trp Leu Leu Leu Ala
            1               5                  10

GGG ACG GTG TTT CCA CAG GGC TGC TGT CAG CAC TGG TCA TAC GGA CTG     157
Gly Thr Val Phe Pro Gln Gly Cys Cys Gln His Trp Ser Tyr Gly Leu
     15                  20                  25

AGC CCA GGA GGG AAG AGG GAT CTG GAC AAC TTC TCA GAC ACA CTG GGC     205
Ser Pro Gly Gly Lys Arg Asp Leu Asp Asn Phe Ser Asp Thr Leu Gly
 30                  35                  40                  45

AAT ATG GTT GAA GAG TTC CCA CGC GTC GAA GCA CCT TGC AGT GTT TTC     253
Asn Met Val Glu Glu Phe Pro Arg Val Glu Ala Pro Cys Ser Val Phe
                 50                  55                  60

GGT TGT GCA GAG GAA TCA CCT TTT GCC AAA ATG TAC AGA GTG AAA GGA     301
Gly Cys Ala Glu Glu Ser Pro Phe Ala Lys Met Tyr Arg Val Lys Gly
             65                  70                  75

CTT CTT GCG AGT GTG GCC GAA AGG AAA ATG GAC ACC GGA CAT TCA AGA     349
Leu Leu Ala Ser Val Ala Glu Arg Lys Met Asp Thr Gly His Ser Arg
         80                  85                  90

AAT GAA AGA TTT CTT TGATTCTACA TTTCATTTTT TATATGAGCA TAAAACATTT     404
Asn Glu Arg Phe Leu
     95

TTTTGTGAAT GTTGCTCTTG TCTTATTATC TAAAATATAA ATAAAAGCTT TCAACTCACT   464

GAAAAAAAAA AAAAAAAAAA AAAAACC                                        491
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Met Ala Ala Lys Ile Leu Ala Leu Trp Leu Leu Leu Ala Gly Thr Val
 1               5                  10                  15

Phe Pro Gln Gly Cys Cys Gln His Trp Ser Tyr Gly Leu Ser Pro Gly
             20                  25                  30

Gly Lys Arg Asp Leu Asp Asn Phe Ser Asp Thr Leu Gly Asn Met Val
         35                  40                  45

Glu Glu Phe Pro Arg Val Glu Ala Pro Cys Ser Val Phe Gly Cys Ala
     50                  55                  60

Glu Glu Ser Pro Phe Ala Lys Met Tyr Arg Val Lys Gly Leu Leu Ala
 65                  70                  75                  80

Ser Val Ala Glu Arg Lys Met Asp Thr Gly His Ser Arg Asn Glu Arg
                 85                  90                  95

Phe Leu
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCACGAATTC CARCAYTGGT CNCA 24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCACGAATTC CARCAYTGGA GYCA 24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCACGAATTC CAYGGNTGGT AYCC 24

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CAGGGCACAC TGTCCTC 17

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCACGAATTC CGAAGCTATG AGCAGTC 27

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCACGAATTC GGTCCCACGG CTGGTAC 27

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCACGAATTC CARCAYTGGT CNTAYGG 27

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCACGAATTC TAYGGNYTNT CNCC 24

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GCACGAATTC TAYGGNYTNA GYCC 24

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CATATTGCCC AGTGTGTC 18

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (primer)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCACGAATTC GTGTCTGAGA AGTTGTCC 28

WHAT IS CLAIMED IS:

1. An isolated nucleic acid encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone.

2. A composition according to claim 1 wherein said nucleic acid encodes a full-length [His^5,Trp^7,Tyr^8]-GnRH preprohormone.

3. A composition according to claim 1 wherein said preprohormone consists of the amino acid sequence depicted in Seq. ID No. 2.

4. A composition according to claim 1 wherein said preprohormone is of fish origin.

5. An isolated nucleic acid encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide.

6. A composition according to claim 5 wherein said nucleic acid encodes a full-length [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide.

7. A composition according to claim 6 wherein said nucleic acid encodes the GAP peptide as shown in Figure 2.

8. An isolated vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone.

9. A composition according to claim 8 wherein said preprohormone is specifically immunoreactive with antibodies raised against an immunogen consisting essentially of a polypeptide of Seq ID No. 2.

10. A composition according to claim 8 wherein said preprohormone is recombinantly produced.

11. A composition according to claim 8 wherein said preprohormone is of fish origin.

12. A composition according to claim 8 wherein said preprohormone is full-length.

13. An isolated vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide.

14. A composition according to claim 13 wherein said GAP peptide is competent to bind antibodies raised against an immunogen consisting essentially of a GAP peptide as shown in Figure 2.

15. A composition according to claim 13 wherein said GAP peptide is recombinantly produced.

16. A composition according to claim 13 wherein said GAP peptide is of fish origin.

17. A composition according to claim 13 wherein said GAP peptide is full-length.

18. An antibody specifically immunoreactive with a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone.

19. A composition according to claim 18 wherein said antibody is specifically immunoreactive with a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone GAP peptide.

20. A nucleic acid probe capable of selectively hybridizing to a nucleic acid encoding a vertebrate GnRH preprohormone.

21. The composition of claim 20 wherein said nucleic acid consists of Seq. ID No. 1.

22. The composition of claim 20 wherein said nucleic acid consists of a GAP peptide cDNA sequence as shown in Seq. ID No. 1.

23. A transgenic non-human animal having increased reproductive capacity due to the expression of a DNA construct encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone introduced into said animal or an ancestor of said animal.

24. The non-human animal of claim 23 wherein said animal is a fish.

25. A method of detecting a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone in a biological sample comprising the steps of:

a) contacting said biological sample with an antibody specifically immunoreactive with a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone;

b) incubating said antibody with said biological sample to form an antibody:[His^5,Trp^7,Tyr^8]-GnRH preprohormone complex; and

c) detecting said complex.

26. A method according to claim 25 wherein said biological sample is of fish origin.

27. A method of detecting a nucleic acid encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone in a biological sample comprising:

a) contacting said biological sample with a nucleic acid probe capable of selectively hybridizing to said nucleic acid encoding a vertebrate [His^5,Trp^7,Tyr^8]-GnRH preprohormone;

b) incubating said nucleic acid probe with the biological sample to form a hybrid of the nucleic acid probe with complementary nucleic acid sequences present in the biological sample; and

c) determining the extent of hybridization of the nucleic acid probe to the complementary nucleic acid sequences.

28. A method according to claim 27 wherein said biological sample is of fish origin.

29. A composition according to claim 1 wherein said preprohormone is of mammalian origin.

30. A composition according to claim 1 wherein said preprohormone consists of the amino acid sequence depicted in Seq. ID No. 18.

31. A composition according to claim 8 wherein said preprohormone is of mammalian origin.

32. A composition according to claim 8 wherein said preprohormone consists of the amino acid sequence depicted in Seq. ID No. 18.

33. An isolated nucleic acid encoding a vertebrate [Ser^8]-GnRH preprohormone.

34. A composition according to claim 33 wherein said preprohormone is of fish origin.

35. A composition according to claim 33 wherein said preprohormone consists of the amino acid sequence depicted in Seq. ID No. 20.

36. An isolated vertebrate [Ser^8]-GnRH preprohormone.

37. A composition according to claim 36 wherein said preprohormone is of fish origin.

38. A composition according to claim 36 wherein said preprohormone consists of the amino acid sequence depicted in Seq. ID No. 20.

Fig. 1:

	GnRH form	Source	1	2	3	4	5	6	7	8	9	10	Sequence ID No.
1a													
	GnRH	Mammal	pE	H	W	S	Y	G	L	R	P	GNH2	4
	GnRH [^{8}Gln]	Chicken (I)	pE	H	W	S	Y	G	L	Q	P	GNH2	5
	GnRH [^{7}Trp^{8}Leu]	Salmon	pE	H	W	S	Y	G	W	L	P	GNH2	6
	GnRH [^{5}His^{8}Asn]	Catfish	pE	H	W	S	H	G	L	N	P	GNH2	7
	GnRH [^{5}His^{7}Trp^{8}Tyr]	Chicken (II)	pE	H	W	S	H	G	W	Y	P	GNH2	8
	GnRH [^{5}His^{7}Trp^{8}Leu]	Dogfish	pE	H	W	S	H	G	W	L	P	GNH2	9
	GnRH [^{3}Tyr^{5}Leu^{6}Glu^{7}Trp^{8}Lys]	Lamprey	pE	H	Y	S	L	E	W	K	P	GNH2	10
	GnRH [^{3}Tyr^{5}His^{6}Asp^{7}Trp^{8}Lys]	Lamprey	pE	H	Y	S	H	D	W	K	P	GNH2	11
		α-yeast mating factor		H	W		L	E	L	K	P	G	12
1b	PCR 5' Chains												
	GnRH (^{7}Trp^{8}Leu)		pE	H	W	S	Y	G	W	L	P	GNH2	
	GnRH (^{5}His^{7}Trp^{8}Tyr)		pE	H	W	S	H	G	W	Y	P	GNH2	
		round 1:	5'-CAA/G	CAT/C	TGG	TCX AGT/C	CA-3'						13
		round 2:					5'-CAT/C	GGX	TGG	TAT/C	CC-3'		14

Fig. 2:

	-20	-10	+1	10
GnRH{^{7}Trp^{8}Leu}:	MEAGSRVIMQVLLLALVVQVTLS **QHWSYGWLPG** GKR			
GnRH{^{5}His^{7}Trp^{8}Tyr}:	MCVSRLALLLGLLLCVGAQLSFA **QHWSHGWYPG** GKR			

20 30 40 50 60

SVGELEATIRMMGTGGVVSLPDEANAQIQERLRPYNIINDDSSHFDRKKRFPNN*

ELDSFGTSEISEEIKLCEAGECSYLRP<u>QRR</u>SILRNILLDALARELQKRK*

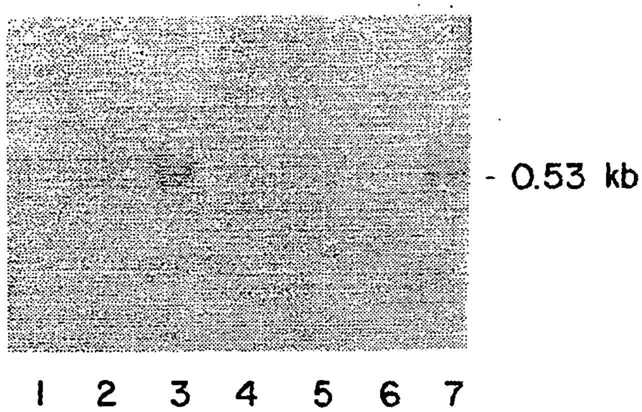


FIG. 3.

Fig. 4:

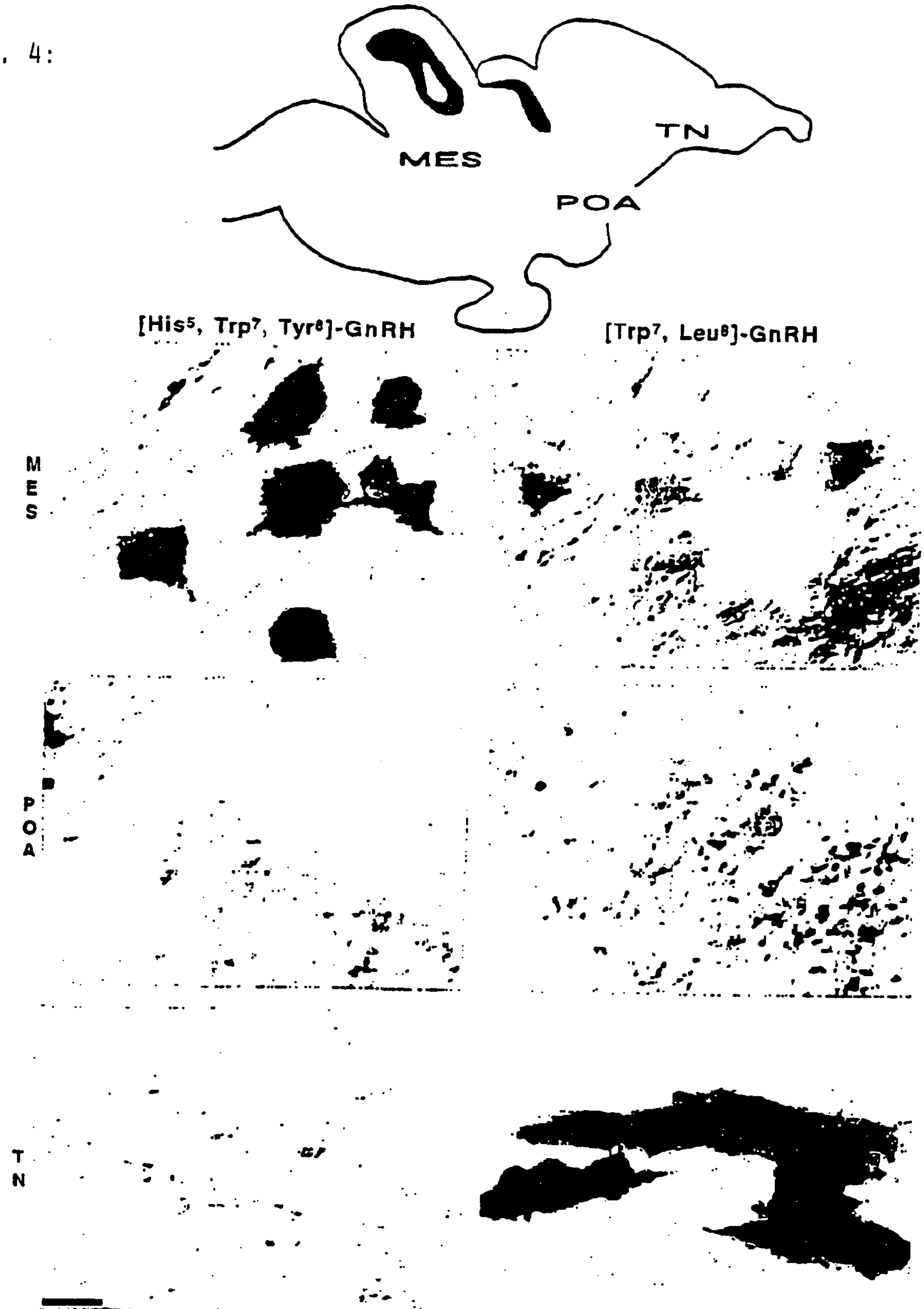
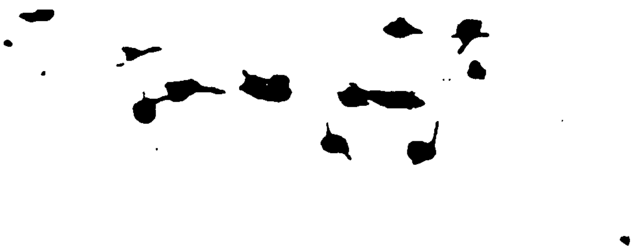


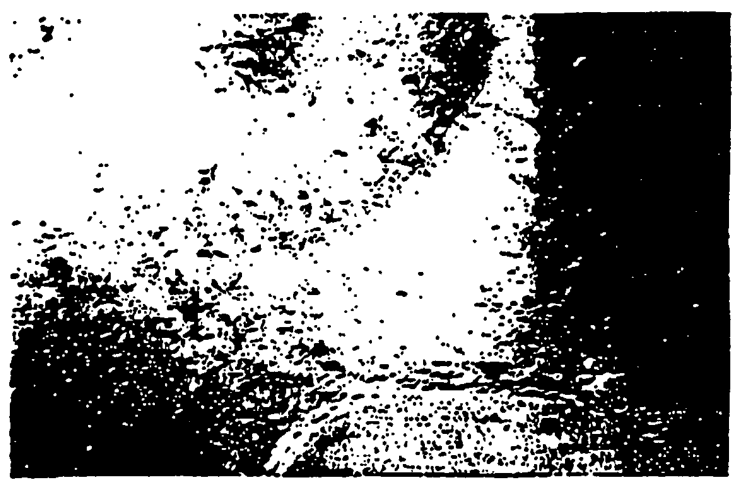
Fig. 5:

PREOPTIC AREA
NON-TERRITORIAL MALE

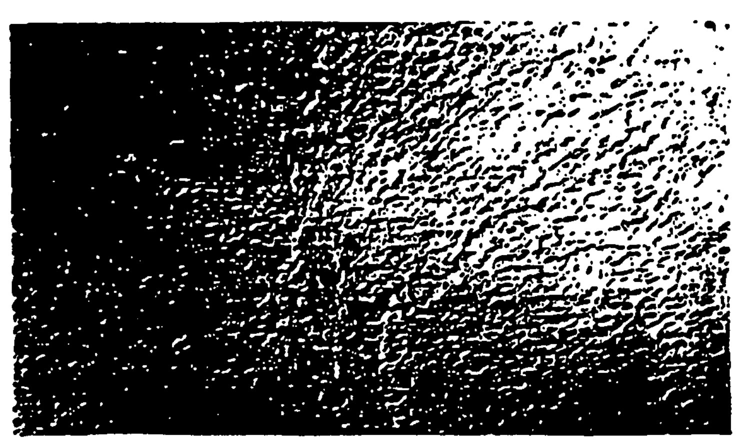


PREOPTIC AREA
TERRITORIAL MALE

TERMINAL NERVE
TERRITORIAL MALE



MESENCEPHALON
TERRITORIAL MALE



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/12763

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/2, DIG1; 536/23.51, 24.31; 530/313, 324, 387.1, 388.1; 435/69.1, 172.3; 436/513

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Molecular Endocrinology, Volume 5, issued 1991, C. Bond et al., "Characterization of Complementary DNA Encoding the Precursor for Gonadotropin-Releasing Hormone and its Associated Peptide from a Teleost Fish", pages 931-937, see entire document.	1-38
A	Society for Neuroscience Abstracts, Volume 19, part 2, issued 1993, S. White et al., "A Second Gene for GNRH: Complementary DNA and Expression Pattern for the Pre-Prohormone in a Teleost Fish", page 1623, see entire abstract.	1-38

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

04 JANUARY 1995

Date of mailing of the international search report

17 FEB 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

CHARLES C. P. RORIES, Ph.D.

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★